

THESE

présentée devant l'UNIVERSITE de NANTES

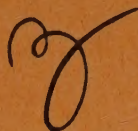
pour obtenir le grade de

DOCTEUR ès-SCIENCES

par

Philippe de LARMINAT

ECOLE NATIONALE SUPERIEURE DE MECANIQUE



SUR L'IDENTIFICATION PAR FILTRAGE NON LINEAIRE

Soutenue le 4 mai 1971 devant la Commission d'Examen

MM. D. TARDY

Président

R. MEZENCEV

M. PELEGRIN

C. MIRA

L. BADGUERAHANIAN

} Examineurs

THESE

présentée devant l'UNIVERSITE de NANTES

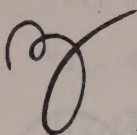
pour obtenir le grade de

DOCTEUR ès-SCIENCES

par

Philippe de LARMINAT

ECOLE NATIONALE SUPERIEURE DE MECANIQUE



SUR L'IDENTIFICATION PAR FILTRAGE NON LINEAIRE

Soutenue le 4 mai 1971 devant la Commission d'Examen

MM. D. TARDY

Président

R. MEZENCEV

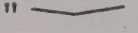
M. PELEGRIN

C. MIRA

L. BADGUERAHANIAN

Examineurs

ERRATA

Page 3, dernière ligne	"... égal à 3 ..."
Page 4, 16ème ligne	"... aplanir ..."
Page 8, 16ème ligne	"... préférentielle ..."
Page 32, 3ème à partir du bas	" $s_{t+1} = [10...0] \underline{x}_{t+1} + \delta'_{t+1} \dots$ "
Page 33, 9ème ligne	" $\underline{z}_k = \underline{x}_t$ "
Page 34, 8ème à partir du bas	" $\underline{z}_{q+1} = f_q(\underline{z}_q, \beta_q)$ "
Page 48, fig. 3-3-b	" $b = - 0,49$ "
Page 53, 6ème ligne	" $\hat{b}_T = - 0,34$ "
Page 59, 14ème ligne	"... $E(\delta_t/s_t) = 0,$ "
Page 60, 5ème ligne	"... $e'_t = e_{t+1} \dots$ "
Page 65, 2ème ligne	" $\overline{\sigma}_t = \hat{\sigma}_t + \frac{2 \hat{\sigma}_t \hat{M}_{\sigma\sigma t}}{D_t} (\dots$ "
Page 65, 7ème ligne	"... $= \hat{\sigma}_{t+1}$ "
Page 68, 11ème ligne	"... $\Phi_t = \sigma_t^2 = (\hat{\sigma}_t + \tilde{\sigma}_t)^2 = \dots$ "
Page 76, 2ème à partir du bas	"... $(\overline{\delta}_t^2 + \overline{M}_{\delta\delta t} - \hat{\Phi}_t)$ "
Page 90, fig. 5-2-3	"  Séquence identifiée directement"
	"  Original de $\hat{H}(z)$ "
Page 93, dans le tableau et dans les trois dernières lignes	Remplacer: $\left \begin{array}{l} a_1 \text{ par } a_0 \\ a_2 \text{ par } a_1 \\ a_3 \text{ par } a_2 \end{array} \right.$
Page 94, dans le tableau	
Page 103, 2ème ligne	" $\dot{x} = \theta_1 + \theta_2 x + \theta_3 u + \theta_5 x^2 + \theta_7 x^3$ "
Page 103, 8ème à partir du bas	"... muni de ces régulations ..."
Page 131, 10ème ligne	" $E(\delta_t / \dots, \underline{x}_{t-1}, \underline{x}_t) = 0$ "
Page 131, 11ème ligne	" $E(\delta'_t / \dots, \underline{x}_{t-1}, \underline{x}_t) = 0$ "

UNIVERSITE DE NANTES

Président M. KERNEIS Médecine

Vice-Président M. PETRESCO Mathématiques

.....

LISTE DES ENSEIGNANTS DES DISCIPLINES SCIENTIFIQUES

Professeurs

M. BRILLOUET

Mathématiques

M. MOLINARO

«

M. PAUC

«

M. PETRESCO

«

M. RAFFIN

«

M. VIVIER (E.N.S.M.)

«

M. TARDY (E.N.S.M.)

Mécanique

M. RAVILLY (E.N.S.M.)

«

M. AVIGNON

Physique

M. BOULLLOUD

«

M. COLLET

«

M. MINN

«

M. MEZENGEV (E.N.S.M.)

«

M. MOUTON

«

M. DUNOYER

Chimie

M. GRAFF

«

M. PIERI

«

M. QUINIOU

«

M. WOJTKOWIAK

«

Melle DEBRAUX

Biologie

M. GARRIGUES

«

Mme LE DOUARIN

«

M. LE DOUARIN

«

M. SELLIER

«

M. CAVET

Sciences de la Terre

SENEQUE : Aie surtout le souci de séparer les choses
du bruit qu'elles font.

Je présente d'abord à nos collègues de l'Association
d'Anatomie de l'E.N.S.M., sous la direction de
Monsieur le Professeur LEBLANC, auquel j'adresse un
vif remerciement pour la collaboration précieuse et les
précieuses nouvelles qu'il m'a fournies tout au long de cette
étude.

Je rends hommage à Monsieur le Professeur LEBLANC,
Monsieur de l'E.N.S.M., qui a bien voulu accepter la tâche
et se dévouer à moi. Je remercie M. LEBLANC de m'avoir
d'abord accepté de faire ce travail.

Que Monsieur LEBLANC me permette de lui adresser
remerciements pour m'avoir permis d'être en contact
personnel avec l'E.N.S.M.

Je tiens aussi à remercier tous ceux qui m'ont
prêté leur aide. Je pense particulièrement à M. FLAHER, avec
moi, à M. BARRON, ABDALAN, ANGER, CHAIX, BIGNARD, et
Mlle TALLET, sans lesquels la partie expérimentale de ce
travail aurait été bien plus limitée.

Je ne saurais oublier non plus les personnes que
je dois à tout de mes premiers professeurs et qu'on ne peut
pas oublier.

Il m'est agréable enfin de rappeler aux collègues
sur lesquels j'ai pu compter au premier plan. A Magdeleine
m'ont été très utiles par le D.E.M., et la présence de
M. RICHAUD et de l'environnement ont bien aidé.

La présente thèse a été entreprise au Laboratoire d'Automatique de l'E.N.S.M., sous la direction de Monsieur le Professeur MEZENCEV, auquel j'exprime ma profonde reconnaissance pour la sollicitude constante et les précieux conseils qu'il m'a prodigués tout au long de cette étude.

Je rends hommage à Monsieur le Professeur TARDY, Directeur de l'E.N.S.M., qui a bien voulu présider le Jury, et je dis ma gratitude à MM. PELEGRIN et BADGUERAHANIAN d'avoir accepté de juger ce travail.

Que Monsieur MIRA trouve ici l'expression de ma reconnaissance pour s'être intéressé à moi en qualité de parrain au C.N.R.S.

Je tiens aussi à remercier tous ceux qui m'ont apporté leur aide. Je pense particulièrement à M. PIASCO, mais aussi à MM. BARRAUD, ABDALAH, AUGIER, COLLIN, BEYSENS, et Melle TALLEC, sans lesquels la partie expérimentale de ce travail aurait été bien plus limitée.

Je ne saurais oublier non plus les ouvertures que je dois à tant de discussions fructueuses et amicales avec M. THOMAS.

Il m'est agréable enfin de rappeler que les données sur lesquelles j'ai vu apparaître mes premiers résultats m'ont été communiquées par le DERA, en la personne de M. RICHALET et je l'en remercie ici bien sincèrement.

SOMMAIRE

	<i>Pages</i>
INTRODUCTION -	1
Chapitre 1 - ASPECTS GENERAUX DE L'IDENTIFICATION -	7
1-1 Formulation - systèmes linéaires discrets - propriétés générales des algorithmes d'identification	7
1-2 Introduction à l'identification par filtrage non linéaire	12
Chapitre 2 - ETUDE DU FILTRAGE NON LINEAIRE -	15
2-1 Equations d'un filtre discret approximativement optimal	15
2-2 Modalités générales d'application du filtrage non linéaire à l'identification	23
2-3 Différents points de vue selon lesquels on peut considérer le même algorithme d'identification	27
Chapitre 3 - MODALITES PRATIQUES DE LA MISE EN OEUVRE DU FILTRAGE NON LINEAIRE -	31
3-1 Décomposition de l'équation d'état généralisée en une suc- cession d'opérations élémentaires	31
3-2 Identification des systèmes continus	37
3-3 Application "hors ligne" et "en ligne". Elimination des connaissances à priori	41
3-4 Etude d'une simulation	43
Chapitre 4 - ESTIMATION DES VARIANCES -	55
4-1 Première approche du problème	55
4-2 Mise en forme telle que les caractéristiques statistiques des bruits soient identifiables par filtrage non linéaire approché	58
4-3 Simplification des équations d'estimation de variance, extension au cas multivariable	67
4-4 Etude d'une simulation	74

Chapitre 5 - RESULTATS EXPERIMENTAUX -	81
5-1 Identification d'un micro-réseau de distribution d'électricité	81
5-2 Identification d'une batterie d'extraction liquide - liquide	87
5-3 Identification de la chaîne de tangage d'un hélicoptère	90
5-4 Identification d'un turbo-réacteur	95
5-5 Identification d'une unité pilote générateur de vapeur	103
Chapitre 6 - APPLICATIONS PARTICULIERES DU FILTRAGE LINEAIRE A L'IDENTIFICATION -	107
6-1 Identification des systèmes linéaires représentés par leur séquence de pondération	108
6-2 Identification par fonctions modulatrices et filtrage linéaire	117
CONCLUSION -	127
Annexe I - LE FILTRE DE KALMAN DISCRET -	131
Annexe II - DESCRIPTION ET LISTING DES SOUS-PROGRAMMES D'IDENTIFICATION -	137
Annexe III - ORGANIGRAMME D'IDENTIFICATION D'UN SYSTEME LINEAIRE RECURRENT DU 2 ^e ORDRE	141
BIBLIOGRAPHIE -	145

I N T R O D U C T I O N

Avant d'aborder l'objet précis de cette étude - l'identification par filtrage non linéaire -, il convient de définir dans quel cadre elle se situe.

On peut dire qu'identifier, c'est rechercher un modèle mathématique qui soit "identique" à un phénomène ou à un objet réel, c'est-à-dire tel que le résultat des calculs concorde avec ceux de l'expérience. Sans discuter de cette identité au niveau de la philosophie des sciences, on peut essayer de distinguer en quoi l'optique de l'automaticien peut s'écarter des points de vue usuels.

Lorsque l'ingénieur effectue une "mise en équation", c'est naturellement en vue d'un objectif déterminé, par exemple : établir les dimensions d'un échangeur de chaleur ou d'un réacteur chimique. En fonction de cet objectif, il sera amené à négliger certains aspects de la réalité, malgré son désir d'établir des équations valables dans le plus vaste domaine de circonstances possibles. L'automaticien, pour réaliser ses propres objectifs, devra généralement donner la première place aux caractéristiques dynamiques de l'ensemble du processus considéré, caractéristiques qui souvent apparaissent mal dans les équations que le spécialiste (chimiste, thermodynamicien ou autre) aura établies primitivement pour des besoins différents. En effet, même si les coefficients issus de l'analyse physico-mathématique sont donnés avec une certaine précision, le modèle dynamique global obtenu en accumulant toutes ces équations (et leurs erreurs !) pourra être beaucoup moins satisfaisant. D'autre part, pour traiter ces équations, l'automaticien sera parfois contraint d'opérer des simplifications, et il est alors indispensable d'en contrôler la validité par une expérimentation globale sur l'ensemble du processus.

Selon cette ligne de conduite, l'analyse, c'est-à-dire la mise en équation de chaque composant ou phénomène élémentaire, tient encore une place prépondérante et fournit au moins qualitativement la forme générale des équations du modèle.

Cependant, et c'est là que l'automaticien s'écartera de la méthode analytique habituelle à l'ingénieur, il cherchera parfois à synthétiser directement un modèle global à partir des données de l'expérience, si besoin est en

adoptant des équations n'ayant plus qu'un rapport lointain avec la "physique" du système telle que l'aurait définie le spécialiste, mais qui soient susceptibles d'être traitées par les méthodes générales de l'automatique (ou théorie de la commande, contrôle de processus, théorie des systèmes, comme on voudra l'appeler).

Ces équations n'auront pas alors pour fonction la connaissance interne du processus, mais de rendre compte de son comportement observable dans le domaine de fonctionnement utile : en d'autres termes, de permettre une prévision de l'évolution des grandeurs à contrôler (sorties) en fonction des commandes *effectivement possibles* (entrées).

A la limite, le processus sera considéré comme une "boîte noire", en exprimant par là que les informations proviennent uniquement de l'observation des signaux d'entrée-sortie du système soumis à des excitations diverses.

Dans la mesure où cette "boîte noire" n'est pas seulement l'image d'un composant élémentaire, mais de tout l'ensemble d'un processus, le modèle mathématique présentera nécessairement un certain degré de complexité, et l'estimation des paramètres des équations devra donc généralement faire appel à des méthodes suffisamment puissantes pour aboutir.

Il se trouve qu'un nombre croissant de ces méthodes d'identification se trouve en rapport étroit avec les théories générales de l'Automatique.

C'est le cas de l'*Identification par filtrage non linéaire*, qui fait l'objet de la majeure partie de notre étude.

A l'origine, un filtre est un procédé de calcul, analogique ou numérique, qui permet de séparer dans un signal observé le "message" des perturbations ou erreurs de mesure.

Cette notion de signal et de message peut se généraliser à tout ensemble de mesure d'où il importe d'extraire la valeur de certaines variables. Dans cette optique, nous considérons les observations sur les entrées et sorties d'un système comme les signaux à partir desquels on doit estimer le message constitué par les paramètres des équations du modèle du processus.

Depuis les premiers travaux de WIENER vers 1942, le filtrage a fait un pas décisif, en 1961, avec la théorie de KALMAN et BUCY, où l'on passe de la représentation externe des processus aléatoires, par fonction de corrélation, à l'exploitation d'une représentation interne sous forme d'équations d'état différentielles ou récurrentes stochastiques.

C'est en 1962 que KOOP et ORFORD ont proposé de considérer les paramètres définissant le modèle mathématique d'un processus comme une partie de

l'état d'un système plus vaste auquel on pourrait appliquer les techniques d'estimation par filtrage. Il ne s'agit pas là d'un simple formalisme destiné à faire le lien entre des procédés établis et une théorie générale, mais véritablement d'un point de vue fécond et qui a permis d'établir de nouveaux algorithmes d'identification de plus en plus efficaces.

Le filtrage non linéaire conduit en effet directement à des algorithmes séquentiels ("en ligne"), susceptibles d'être appliqués à l'identification en temps réel de paramètres non stationnaires.

Il apparaît cependant que, même pour des systèmes linéaires, les équations d'état généralisées que l'on doit considérer pour l'identification sont, sauf cas exceptionnels, non linéaires.

Les premiers travaux ont justement été réalisés dans ces cas exceptionnels : celui, par exemple, de mesures linéaires par rapport aux paramètres à identifier : Cf. R.C.K. LEE (1964), FARISON (1967). Ces méthodes, qui consistent alors en une formulation récurrente des moindres carrés ordinaires, sont malheureusement biaisées (et même asymptotiquement biaisées) lorsque les mesures sont entachées de bruit.

Pour dépasser ce stade, une autre approche a été largement utilisée : elle consiste à linéariser les non linéarités autour d'un développement limité au premier ordre et à établir un filtre de KALMAN non stationnaire à partir des équations ainsi obtenues : BRYSON (1963), COX (1964), MOWERY (1965), SCHWEPPE (1964), KUMAR (1964).

Parallèlement, le filtrage non linéaire (indépendamment de ses applications possibles à l'identification) était progressivement résolu par STRATONOVITCH (1960), KUSHNER (1964), WONHAM (1965), BUCY (1968) et autres.

Les équations exactes du filtrage non linéaire sont malheureusement des équations aux dérivées partielles, dont la résolution analytique ou numérique est pratiquement toujours irréalisable (sauf dans le cas linéaire où l'on retrouve le filtre de KALMAN-BUCY).

C'est pourquoi les auteurs ci-dessus ont recherché également des solutions approchées en faisant certaines approximations : celles-ci peuvent porter sur le processus lui-même, dont les non linéarités sont approchées par des développements en série de TAYLOR dans lesquels on retient cette fois les termes jusqu'au second ordre. On doit faire également des hypothèses sur les distributions de probabilités conditionnelles : les plus communes consistent soit à faire l'hypothèse gaussienne, soit à supposer que les moments d'ordre supérieur ou égal à 2 sont négligeables.

Selon la nature exacte des hypothèses ci-dessus, selon aussi le problème considéré (processus continu avec observations discrètes, ou purement continu, ou bien purement discret...), un certain nombre d'algorithmes ont été proposés par divers auteurs. Citons : JAZVINSKI (1966), ATHANS (1968), NEAL (1968), BUDIN (1969), FAROOK (1970)...

La supériorité de ces méthodes (du 2ème ordre) semble maintenant établie sur le filtrage de KALMAN appliqué à une simple linéarisation du premier ordre, mais il ne semble pas que des comparaisons entre elles soient parvenues à des conclusions définitives (cf. SCHWARTZ (1967), ATHANS (1968), TAIT (1970)).

Ces comparaisons portent d'ailleurs presque toujours sur des simulations et c'est récemment que des applications à l'identification de processus réels se sont fait connaître : WELLS (1969), SASTRY (1969), HABEGGER (1969).

Le nombre croissant de publications dans le domaine de l'identification par filtrage non linéaire témoigne de l'intérêt de cette méthode, mais nous sommes loin cependant de nous trouver en présence d'une théorie achevée.

Le travail présenté ici a pour objet de contribuer à applanir les difficultés qui s'opposent à la généralisation de son emploi. Nous en avons divisé l'exposé en six chapitres :

Le premier chapitre : *Aspects généraux de l'identification* nous permet d'introduire le principe de l'identification par filtrage non linéaire dans le cadre des propriétés souhaitables pour toute méthode d'identification.

Nous abordons ensuite *l'Etude du filtrage non linéaire discret*. Nous nous sommes limités au filtrage discret en considérant que, le nombre des équations d'estimation croissant comme le carré de la dimension de l'état à estimer, une résolution par voie analogique devient rapidement impraticable. Il apparaît donc indiqué de poser d'emblée le problème en terme discrets plutôt que d'opérer cette discrétisation à posteriori pour les besoins d'une intégration numérique des équations différentielles.

Les équations générales d'un filtre discret approximativement optimal, que nous développons sous une forme synthétique dans ce deuxième chapitre, reposent essentiellement sur l'approximation gaussienne pour les densités de probabilité des estimations successives.

Le troisième chapitre est consacré aux *modalités pratiques de la mise en oeuvre du filtrage non linéaire*. Nous proposons en particulier un ensemble de sous programmes constituant des "macro-instructions" d'identification, permettant de traduire directement les équations d'état généralisées en un programme d'identification modulaire sans qu'il soit besoin de rentrer dans le détail des opérations de filtrage.

Nous présentons ensuite (Chapitre 4), une solution originale au problème de *l'estimation des variances* des bruits et perturbations intervenant dans le processus, ce qui permet d'établir un filtre d'estimation sans qu'il soit besoin de connaître précisément les caractéristiques statistiques de ces bruits.

Les *Résultats expérimentaux* obtenus par filtrage non linéaire font l'objet du cinquième chapitre. C'est sur eux en effet qu'on doit juger en dernier ressort de la valeur d'une méthode d'identification, lorsque des divergences incontrôlables entre la réalité et le cadre mathématique adopté viennent s'ajouter aux approximations de la méthode elle-même.

Nous présentons donc les résultats obtenus (pour certains en temps réel) sur les problèmes suivants : Micro-réseau de distribution d'électricité, processus chimique, chaîne de tangage d'un hélicoptère, turboréacteur, générateur de vapeur pilote du Laboratoire de l'E.N.S.M.

Deux applications particulières du filtrage linéaire à l'identification sont proposées dans le dernier chapitre. Nous avons dit qu'en général l'identification ne pouvait se réduire à un problème linéaire. C'est le cas cependant lorsqu'on adopte pour les systèmes une représentation par réponse impulsionnelle, et nous proposons une méthode d'initialisation de l'algorithme permettant d'améliorer sensiblement les résultats classiques (LEVIN, 1960).

Dans ce même chapitre, nous rapprochons ensuite le point de vue de l'identification par fonctions modulatrices (LOEB, 1967) de celui de l'identification par filtrage linéaire. Nous obtenons ainsi un algorithme susceptible de conduire à une mise en oeuvre analogique pour l'identification en temps réel de systèmes non stationnaires. Dans le cas d'un traitement digital, nous utilisons les mêmes techniques que pour le filtrage non linéaire, dont nous sommes ici en présence d'un cas particulier.

Il n'est pas question de prétendre que le point de vue du filtrage puisse apporter une solution à l'identification en général. Il nous a permis cependant de traiter des problèmes réels et variés, linéaires ou non, stationnaires ou non, en ligne ou hors ligne, identification de variance ; et ceci non seulement à partir de la même théorie de base, mais encore en utilisant dans tous les cas les mêmes procédés matériels (sous programmes spécialisés élémentaires), c'est-à-dire avec un effort d'adaptation minimum pour chaque cas particulier.

ASPECTS GENERAUX DE L'IDENTIFICATION

1-1 Formulation - Systèmes linéaires discrets -

Propriétés générales des algorithmes d'identification

1-2 Introduction à l'identification par filtrage non linéaire

Dans ce chapitre, nous exposons un problème type d'identification, celui des systèmes linéaires discrets. Cet exemple nous permet de discuter concrètement des propriétés généralement souhaitables pour un algorithme d'identification, de donner un premier aperçu de l'identification par filtrage non linéaire, et de situer cette solution par rapport à d'autres méthodes.

1-1 FORMULATION - SYSTEMES LINEAIRES DISCRETS - PROPRIETES GENERALES DES ALGORITHMES D'IDENTIFICATION -

Afin d'introduire concrètement le point de vue choisi parmi les différentes façons d'aborder le problème de l'identification, nous prenons comme point de départ l'identification des systèmes dynamiques linéaires discrets.

Parmi ceux-ci, on retient ceux qui peuvent être représentés par des équations d'état de la forme suivante :

$$\underline{y}_{t+1} = A\underline{y}_t + B\underline{u}_t \quad 1-1-1$$

$$\underline{s}_{t+1} = C\underline{y}_{t+1} \quad 1-1-2$$

$t = 0, 1, 2, \dots, T$ (indice temporel)

$\underline{y}_t$: vecteur d'état (dimension n , la notation $\underline{x}_t$ sera réservée à l'état d'un processus généralisé incluant les paramètres à identifier)

$\underline{s}_t$: séquence de sortie (dimension r)

$\underline{u}_t$: séquence d'entrée (dimension p)

A, B, C : matrices ($n \times n$), ($n \times p$), et ($r \times n$)

La figure (1-1) symbolise ces équations.



fig. 1-1

En première approximation, on peut définir l'identification comme la recherche de la valeur des paramètres du système (ici les coefficients des matrices A, B, C), à partir de la connaissance approchée (en raison des perturbations et erreurs de mesure) des séquences d'entrée et de sortie $\underline{u}_t$, $\underline{y}_t$ ($0 < t < T$).

On admet ici implicitement que les coefficients à identifier "existent", ce qui n'aura plus de sens quand on voudra identifier un système complexe par un modèle simplifié, et nous verrons alors à utiliser une définition plus générale.

Le problème ci-dessus, pourtant relativement restreint, a fait l'objet jusqu'ici d'un nombre considérable d'études. Les méthodes proposées sont extrêmement variées et, si l'on peut distinguer quelques directions préférencielles, il ne semble pas qu'une méthode particulière puisse être considérée comme fondamentale, au point de constituer une théorie de base comme il peut en exister dans le domaine de la commande optimale de processus linéaire.

A cela, deux explications peuvent être proposées :

1°) Le problème ci-dessus n'est pas complètement défini, en particulier en ce qui concerne :

- les erreurs de mesure
- le type des signaux d'entrée
- la définition de l'équivalence entre le processus réel et le modèle adopté (dans le cas où ils sont de structure différente)
- la forme particulière des matrices A, B, C

Selon les précisions apportées, il peut découler des problèmes distincts exigeant des solutions différentes.

2°) Il serait cependant possible de choisir un cadre relativement général et réaliste, dans lequel la majorité des problèmes réels puisse s'insérer (en se limitant éventuellement aux systèmes linéaires).

Si nous considérons alors les principales propriétés souhaitables pour un algorithme d'identification (exigences qui à priori ne paraissent pas mutuellement contradictoires), nous pourrions constater qu'il n'existe pas de méthode générale qui réunisse toutes les qualités requises sur le plan théorique et pratique. Ce fait explique l'abondance des méthodes différentes, selon que l'accent est mis sur une propriété ou sur une autre.

Nous avons dit qu'un algorithme d'identification opère sur des données qui sont les séquences d'entrées - sorties (u_t, y_t) d'un processus à identifier, afin de fournir une estimation des paramètres qui définissent les équations de ce processus.

En se limitant au cas des systèmes linéaires discrets, nous pouvons énumérer quelques propriétés essentielles qu'un tel algorithme devrait présenter pour être considéré comme satisfaisant d'une manière générale :

- A - Ne pas exiger un volume de calcul et de mémoire démesuré.
- B - Converger vers les valeurs exactes des paramètres quand le nombre de mesures disponibles s'accroît.
- C - Fournir des résultats d'une précision au moins voisine du maximum théoriquement atteignable.
- D - Donner une estimation de la précision des résultats.
- E - S'il s'agit d'un algorithme récurrent, converger quel que soit le point de départ choisi.
- F - Etre généralisable à des systèmes multivariables, d'ordre élevé, ou non stationnaires.
- G - Conserver des propriétés convenables même si les bruits s'écartent sensiblement de leurs caractéristiques présumées.
- H - Etre adapté à une large classe de signaux d'entrée.

Nous avons avancé qu'il n'existe pas de méthode qui satisfasse toutes ces conditions simultanément. Nous revenons donc sur certaines d'entre elles.

(A) : le volume de calcul. Malgré le développement des calculateurs modernes, on peut difficilement généraliser l'usage de méthodes demandant l'intégration numérique d'équations aux dérivées partielles (filtrage non linéaire exact), l'exploration systématique de l'espace des paramètres

(dans l'identification par recherche d'un minimum), l'utilisation intermédiaire d'une représentation non paramétrique (matrice des réponses impulsionnelles), etc...

Pour discuter des points (B),(C),(D), il est nécessaire de préciser le type des erreurs de mesure.

A titre d'exemple, on revient aux systèmes linéaires discrets : les équations (1-1-1) et (1-1-2) sont alors précisées comme suit :

$$\underline{y}_{t+1} = A\underline{y}_t + B\underline{u}_t + \underline{\beta}_t \quad 1-1-3$$

$$\underline{s}_{t+1} = C\underline{y}_{t+1} + \underline{\gamma}_{t+1} \quad 1-1-4$$

$\underline{\beta}_t$ est une séquence perturbation, ou erreur de mesure sur les entrées.

$\underline{\gamma}_t$ est une séquence d'erreur de mesure sur la sortie.

Les données de l'identification restent les séquences $(\underline{u}_t, \underline{s}_t)$.

L'hypothèse la plus commune consiste à supposer que les séquences $\underline{\beta}$ et $\underline{\gamma}$ sont gaussiennes, indépendantes, et de moyenne nulle.

Etant donné un processus, il est alors théoriquement possible de calculer les propriétés statistiques d'un algorithme d'estimation donné quelconque.

(B) On constate alors que certaines méthodes, telles que "Minimisation d'erreur d'équation", présentent une erreur systématique (biais):cf. Astrom, 1970, p. 17.

(C) Il est possible de définir d'un point de vue statistique ce qu'est l'optimalité pour un algorithme d'identification : *variance minimale* par exemple.

Bien qu'un tel critère ne soit pas universel, il est impossible de considérer comme généralement satisfaisant un algorithme dont la précision se situerait très en dessous de l'optimum théorique, ce qui est le cas de méthodes ne prenant en considération que certains points caractéristiques de réponses type (indicielle ou impulsionnelle), ou de méthodes axées essentiellement sur la stabilité ou la simplicité de calcul. Celles-ci gardent cependant leur intérêt dans leur domaine d'application, soit que des résultats approximatifs soient suffisants, soit que les mesures soient extrêmement précises ou surabondantes.

Il est à noter que, pour le modèle linéaire discret, avec perturbations gaussiennes, défini ci-dessus (1-1-3, 1-1-4), modèle qui présente

cependant la simplicité maximale compatible avec une certaine généralité, *il n'existe pas d'algorithme fini de variance minimale* (sauf perturbations de sortie nulles).

Il existe cependant un arsenal de méthodes dont le voisinage de l'optimalité est démontré ou au moins présumé.

Citons :

- le maximum de vraisemblance (M. V.) (Åström et Bohlin, 1965)
- la méthode du modèle (Richalet, 1966) : méthode qui, sous certaines hypothèses, est équivalente au M.V.
- les moindres carrés répétés ou généralisés (Clarke, 1967)

La méthode du maximum de vraisemblance, par exemple, permet de calculer une estimation de la précision des estimations (D). Cependant, elle se traduit par la recherche d'un minimum, et, selon le type du processus et des signaux d'entrée, les algorithmes de programmation non linéaire pourront converger plus ou moins rapidement, ou même diverger (E) si le point de départ est mal choisi.

Par ailleurs, ces méthodes sont le plus souvent difficilement généralisables, au moins sur le plan pratique, à des systèmes multivariables ou de dimension élevée (F).

Tous ces algorithmes exigent d'autre part d'effectuer un certain nombre d'itérations sur la totalité de l'enregistrement d'entrée - sortie ($\underline{u}_t, \underline{y}_t : 0 < t < T$), ce sont donc des méthodes spécifiquement "hors ligne".

Les difficultés de l'identification "hors ligne" (c'est-à-dire à partir d'un enregistrement disponible) se retrouvent évidemment dans l'identification "en ligne", définie par la contrainte supplémentaire qu'on ne peut garder en mémoire l'ensemble de toutes les mesures passées depuis l'instant initial jusqu'à l'instant présent. On peut donc affirmer que, là aussi, la solution idéale n'existe pas, ni pour le problème réduit (linéaire, discret, gaussien), et encore moins pour l'identification en général.

Pour illustrer ce point de vue, nous n'avons cité que quelques méthodes. On pourra trouver un tableau plus complet de l'état actuel de l'identification dans "System Identification" par K.J. Åström et P. Eykhoff (Survey Paper, 2ème I.F.A.C. Symposium on Identification, Prague 1970).

1-2 INTRODUCTION A L'IDENTIFICATION PAR FILTRAGE NON LINEAIRE -

En l'absence d'une méthode générale satisfaisante, toute solution sera nécessairement un compromis entre les exigences énumérées au paragraphe précédent.

Pour notre part, nous avons mis l'accent sur le point (C) : fournir des résultats d'une précision au moins voisine du maximum théoriquement atteignable, en adoptant le principe de l'identification par filtrage non linéaire :

Nous reprenons le système d'équations (1-1-3, 1-1-4) avec l'hypothèse que les séquences aléatoires β et γ sont gaussiennes, indépendantes et centrées :

$$\underline{y}_{t+1} = A\underline{y}_t + B\underline{u}_t + \underline{\beta}_t \quad 1-1-3$$

$$\underline{s}_{t+1} = C\underline{y}_{t+1} + \underline{\gamma}_{t+1} \quad 1-1-4$$

On note $\underline{\theta}$ l'ensemble des paramètres à identifier, coefficients des matrices A, B, C.

On peut définir l'estimation optimale $\underline{\theta}$ à l'instant t :

$$\hat{\underline{\theta}}_t = E(\underline{\theta} | t)$$

où $E(\underline{\theta} | t)$ représente l'espérance de $\underline{\theta}$ conditionnellement aux observations $\underline{s}_k$ pour $k < t$

$\hat{\underline{\theta}}_t$ est alors une *estimation de Bayes* des paramètres, c'est-à-dire qu'elle minimise l'espérance du coût d'une erreur d'identification pour toute fonction de coût quadratique.

Pour inclure le problème de l'identification dans la théorie du filtrage, on introduit une équation complémentaire :

$$\underline{\theta}_{t+1} = \underline{\theta}_t = \underline{\theta} \quad 1-2-6$$

D'où, en notant $\underline{x}_t = \begin{bmatrix} \underline{y}_t \\ \underline{\theta}_t \end{bmatrix}$, vecteur d'état généralisé, on peut réécrire le système (1-1-3, 1-1-4, 1-2-6) sous la forme :

$$\underline{x}_{t+1} = \underline{f}_t(\underline{x}_t, \underline{\beta}_t) \quad 1-2-7$$

$$\underline{s}_{t+1} = \underline{g}_{t+1}(\underline{x}_{t+1}, \underline{\gamma}_{t+1}) \quad 1-2-8$$

L'identification peut être considérée alors comme un sous produit de l'estimation de l'état généralisé $\underline{x}_t$:

$$\hat{\underline{x}}_t = E(\underline{x}_t | t)$$

Il s'agit de déterminer le filtre susceptible de réaliser cette opération.

Le filtrage est un problème parfaitement résolu (Kalman, 1961) si le système (1-2-7, 1-2-8) est linéaire. Par contre, les équations ci-dessus sont généralement non linéaires en $\underline{x}$, d'où la nécessité d'appliquer des techniques de filtrage non linéaire. Notons qu'il n'y a plus de raison alors de se restreindre aux processus linéaires, puisque le système des équations généralisées sera de toutes façons non linéaire.

L'hypothèse stationnaire pourra également être abandonnée en supposant un modèle stochastique (processus de Markov) pour l'évolution des paramètres.

On ne se limitera pas non plus aux systèmes discrets, les processus continus susceptibles d'être représentés par des équations d'état différentielles stochastiques pourront également être identifiés par le même principe du filtrage non linéaire.

Sans recouvrir entièrement le champ de l'identification, ce point de vue apparaît donc très général, d'autant plus qu'il se présente d'emblée comme une identification "en ligne".

En contre-partie, le filtrage non linéaire *optimal* ($\hat{\underline{x}}_t = E(\underline{x}_t | t)$) implique la résolution d'équations aux dérivées partielles (dans le cas continu) ou d'intégrales (dans le cas discret) pour lesquelles il n'existe pas de solution explicite, et leur résolution numérique demande un volume de calcul et de mémoire qui deviendrait rapidement prohibitif, même pour les problèmes les plus simples.

On ne peut donc mettre en oeuvre que des filtres approximativement optimaux.

La difficulté consiste à assurer la *convergence* du filtre en conservant à celui-ci un *caractère quasi optimal*. Sans que ce soit théoriquement incompatible, il n'existe pas à notre connaissance de méthode permettant d'atteindre simultanément ces deux objectifs, sauf cas particulier d'intérêt limité.

Les conditions suffisantes de stabilité (ou même d'hyper-stabilité : Landau, 1969) qu'il est possible d'établir sont en général trop contraignantes

pour être le point de départ de l'élaboration d'un algorithme quasi-optimal.

Pour notre part, nous avons donné priorité à l'optimalité (au sens défini plus haut), en remarquant que si la stabilité n'est pas théoriquement garantie, elle se trouve pratiquement observée dans un grand nombre d'applications concrètes.

Chapitre 2

ETUDE DU FILTRAGE NON LINEAIRE

- 2-1 Equations d'un filtre discret approximativement optimal.
- 2-2 Modalités générales d'application du F.N.L. à l'identification.
- 2-3 Différents points de vue selon lesquels on peut considérer le même algorithme d'identification.

Ce chapitre est consacré aux équations générales d'un filtre non linéaire approché, dans le cas le plus général d'une équation d'état discrète :

$$\underline{x}_{t+1} = \underline{f}_t(\underline{x}_t, \underline{\delta}_t)$$

où $\underline{\delta}_t$ est une séquence aléatoire indépendante centrée.

On montre ensuite comment ces équations peuvent être appliquées à l'identification des processus discrets.

Dans la troisième partie de ce chapitre, nous étudions l'équivalence de l'algorithme obtenu avec d'autres méthodes d'identification, ce qui permet d'envisager son application à des problèmes où ne se trouvent pas réunies les hypothèses de base qui sont :

- Identité des structures du processus et du modèle
- Perturbations susceptibles d'une description statistique.

2-1 EQUATIONS D'UN FILTRE DISCRET APPROXIMATIVEMENT OPTIMAL -

2-1-1 Forme adoptée pour les équations d'état

Nous développerons les équations de filtrage pour un système dont les équations auront la forme suivante :

$$\underline{x}_{t+1} = \underline{f}_t(\underline{x}_t, \underline{\delta}_t) \quad 2-1-1$$

$$\underline{s}_{t+1} = \underline{C}_{t+1} \underline{x}_{t+1} + \underline{\delta}'_{t+1} \quad 2-1-2$$

- $\underline{x}$: vecteur d'état généralisé (dimension n) pouvant inclure des paramètres à identifier.
- $\underline{s}$: vecteur observations (dimension r), fonction *linéaire* de $\underline{x}$
- $\underline{\delta}_t$ et $\underline{\delta}'_t$: séquences indépendantes de $\underline{x}_t, \underline{x}_{t-1}, \dots$, centrées, de covariances :

$$E (\underline{\delta}_\tau \underline{\delta}_t^T) = \begin{cases} \Phi_t & \text{si } \tau = t \\ 0 & \text{si } \tau \neq t \end{cases}$$

$$E (\underline{\delta}'_\tau \underline{\delta}_t^T) = 0$$

$$E (\underline{\delta}'_\tau \underline{\delta}'_t^T) = \begin{cases} \Phi'_t & \text{si } \tau = t \\ 0 & \text{si } \tau \neq t \end{cases}$$

Les variances Φ_t et Φ'_t seront ici des données du problème.

- C_t : matrice $n \times r$
- $\underline{f}_t ()$: fonction supposée deux fois dérivable par rapport à $\underline{x}_t$ et $\underline{\delta}_t$

2-1-2 Domaine d'application

Considérons la forme plus générale

$$\underline{x}_{t+1} = \underline{f}_t(\underline{x}_t, \underline{\beta}_t) \quad 2-1-3$$

$$\underline{s}_{t+1} = \underline{g}_{t+1}(\underline{x}_{t+1}, \underline{\gamma}_{t+1}) \quad 2-1-4$$

$\underline{f}$ et $\underline{g}$ deux fois dérivables, $\underline{\beta}_t$ et $\underline{\gamma}_t$: séquences aléatoires indépendantes centrées :

$$E \left(\begin{bmatrix} \underline{\beta}_t \\ \underline{\gamma}_t \end{bmatrix} \begin{bmatrix} \underline{\beta}_t \\ \underline{\gamma}_t \end{bmatrix}^T \right) = R_t$$

Montrons que ces équations (2-1-3, 2-1-4) peuvent se ramener aux précédentes (2-1-1, 2-1-2)

On introduit les notations suivantes :

$$\underline{y}_t = \begin{bmatrix} \underline{\beta}_t \\ \underline{\gamma}_t \end{bmatrix}, \quad \underline{\delta}_t = \begin{bmatrix} \underline{\beta}_{t+1} \\ \underline{\gamma}_{t+1} \end{bmatrix}, \quad \Phi_t = R_{t+1}$$

Définissons : $\underline{f}'_t(\underline{x}_t, \underline{y}_t) = \underline{f}_t(\underline{x}_t, \underline{\beta}_t)$

$$\underline{g}'_t(\underline{x}_t, \underline{y}_t) = \underline{g}_t(\underline{x}_t, \underline{\gamma}_t)$$

Nous pouvons introduire le système ayant pour état $\begin{bmatrix} \underline{x}_t \\ \underline{y}_t \end{bmatrix}$, soit :

$$\begin{cases} \underline{x}_{t+1} = \underline{f}'_t(\underline{x}_t, \underline{y}_t) \\ \underline{y}_{t+1} = \underline{\delta}_t \end{cases}$$

$$\underline{s}_{t+1} = \underline{g}'(\underline{x}_{t+1}, \underline{y}_{t+1})$$

Ce système est équivalent au système primitif, en ce sens que $\underline{s}_t$ est identique dans les deux cas, et $\underline{\delta}_t$ est une séquence indépendante centrée de matrice de variance-covariance Φ_t connue.

Introduisons un troisième état $\underline{z}$ tel que :

$$\begin{aligned} \underline{z}_{t+1} &= \underline{g}'_{t+1}(\underline{x}_{t+1}, \underline{y}_{t+1}) = \underline{g}'_{t+1}(\underline{f}'_t(\underline{x}_t, \underline{y}_t), \underline{y}_{t+1}) \\ &\triangleq \underline{g}''_t(\underline{x}_t, \underline{y}_t, \underline{\delta}_t) \end{aligned}$$

D'où le système :

$$\begin{cases} \underline{x}_{t+1} = \underline{f}'_t(\underline{x}_t, \underline{y}_t) \\ \underline{y}_{t+1} = \underline{\delta}_t \\ \underline{z}_{t+1} = \underline{g}''_t(\underline{x}_t, \underline{y}_t, \underline{\delta}_t) \end{cases}$$

$$\underline{s}_{t+1} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \underline{x}_{t+1} \\ \underline{y}_{t+1} \\ \underline{z}_{t+1} \end{bmatrix}$$

Ce dernier système équivalent est de la forme générale (2-1-1, 2-1-2) (sortie fonction linéaire de l'état).

Il est même d'une forme plus réduite encore :

$$\underline{x}_{t+1} = \underline{f}_t(\underline{x}_t, \underline{\delta}_t) \quad 2-1-5$$

$$\underline{s}_{t+1} = \underline{C}_{t+1} \underline{x}_{t+1} \quad 2-1-6$$

Nous pourrions donc utiliser sans perte de généralité cette dernière forme. Nous le ferons lorsque se posera le problème de l'estimation des variances des perturbations, où Φ_t ne sera pas une donnée du problème.

2-1-3 Principe du filtrage

Nous considérons des équations

$$\underline{x}_{t+1} = \underline{f}_t(\underline{x}_t, \underline{\delta}_t)$$

2-1-1

$$\underline{s}_{t+1} = C_{t+1} \underline{x}_{t+1} + \underline{\delta}'_{t+1}$$

2-1-2

On désire calculer l'espérance de $\underline{x}$ conditionnellement aux mesures, ou du moins en obtenir une approximation.

La méthode adoptée est la suivante :

- 1°) On remplace l'équation d'état (2-1-1) par une équation approchée :

$$\underline{x}_{t+1} \sim A_t \underline{x}_t + \underline{u}_t + \underline{b}_t$$

2-1-7

A_t : matrice $n \times n$

$\underline{u}_t$: vecteur $n \times 1$

$\underline{b}_t$: vecteur aléatoire indépendant de $\underline{x}_t$, centré, de matrice de covariance Φ_{bt}

Un problème sera la détermination des grandeurs A_t , $\underline{u}_t$, Φ_{bt} . Nous l'étudierons dans la section suivante (2-1-4).

- 2°) A partir des équations 2-1-7 et 2-1-2, on met en oeuvre un filtre de Kalman. Nous en rappelons les équations : (voir annexe 1)

Equations de "prédiction"

$$\underline{x}_{t+1}^* = A_t \hat{\underline{x}}_t + \underline{u}_t$$

2-1-8

$$M_{t+1}^* = A_t \hat{M}_t A_t^T + \Phi_{bt}$$

2-1-9

Equations d' "estimation"

$$\hat{\underline{x}}_{t+1} = \underline{x}_{t+1}^* + M_{t+1}^* C_{t+1}^T (C_{t+1} M_{t+1}^* C_{t+1}^T + \Phi')^{-1} (\underline{s}_{t+1} - C_{t+1} \underline{x}_{t+1}^*)$$

2-1-10

$$\hat{M}_{t+1} = M_{t+1}^* - M_{t+1}^* C_{t+1}^T (C_{t+1} M_{t+1}^* C_{t+1}^T + \Phi')^{-1} C_{t+1} M_{t+1}^*$$

2-1-11

Signification des grandeurs $\underline{x}^*$ et $\hat{\underline{x}}$:

En supposant - L'équation 2-1-7 exacte, avec $\underline{b}_t$ gaussien

- La probabilité de l'état initial $\underline{x}_0$ gaussienne (espérance $\hat{\underline{x}}_0$, covariance $\hat{M}_0$)

- $\underline{\delta}_t$ gaussien

Alors on aurait :

$$\hat{\underline{x}}_t = E (\underline{x}_t / t)$$

$$\hat{M}_t = E ((\underline{x}_t - \hat{\underline{x}}_t) (\underline{x}_t - \hat{\underline{x}}_t)^T / t)$$

et :

$$\underline{x}_t^* = E (\underline{x}_t / t-1)$$

$$M_t^* = E ((\underline{x}_t - \underline{x}_t^*) (\underline{x}_t - \underline{x}_t^*)^T / t-1)$$

Du fait de l'approximation de 2-1-1 par l'équation 2-1-7, $\hat{\underline{x}}$, $\underline{x}^*$, $\hat{M}$, M^* ne seront que des approximations des espérances conditionnelles définies ci-dessus.

Les équations 2-1-8 à 2-1-11 définissent donc un filtre récurrent approximativement optimal pour l'estimation de l'état $\underline{x}_t$;

$$\hat{\underline{x}}_t \sim E (\underline{x}_t / t)$$

Les grandeurs A_t , $\underline{u}_t$, Φ_{bt} interviennent dans les équations de prédiction. La détermination de ces grandeurs pourra donc être faite en utilisant les résultats $\hat{\underline{x}}_t$, $\hat{M}_t$ de l'étape d'estimation précédente.

2-1-4 Détermination de l'approximation linéaire

Nous recherchons les valeurs A_t , $\underline{u}_t$, Φ_{bt} , tels que :

$$\underline{x}_{t+1} = f_t(\underline{x}_t, \delta_t) \sim A_t \underline{x}_t + \underline{u}_t + b_t$$

On introduit les notations suivantes :

(les indices inférieurs sont des indices "temps", les indices supérieurs sont des indices de "composante")

$$\underline{\zeta}_t \triangleq \begin{bmatrix} \underline{x}_t \\ \delta_t \end{bmatrix}, \quad f_t(\underline{\zeta}_t) \triangleq f_t(\underline{x}_t, \delta_t) = \begin{bmatrix} f_t^1(\underline{\zeta}_t) \\ \vdots \\ f_t^n(\underline{\zeta}_t) \end{bmatrix}$$

$$\hat{F}_{\zeta t} \triangleq \begin{bmatrix} \frac{\partial f_t^1}{\partial \zeta^1} & \frac{\partial f_t^1}{\partial \zeta^2} & \dots \\ \frac{\partial f_t^2}{\partial \zeta^1} & \frac{\partial f_t^2}{\partial \zeta^2} & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} \quad \triangleq \begin{bmatrix} \hat{F}_{xt} & \hat{F}_{\delta t} \end{bmatrix} \quad \zeta = \hat{\zeta}_t$$

$$\hat{F}_{\zeta\zeta t}^i \triangleq \begin{bmatrix} \frac{\partial^2 f_t^i}{\partial \zeta^1 \partial \zeta^1} & \frac{\partial^2 f_t^i}{\partial \zeta^1 \partial \zeta^2} & \dots \\ \frac{\partial^2 f_t^i}{\partial \zeta^2 \partial \zeta^1} & \frac{\partial^2 f_t^i}{\partial \zeta^2 \partial \zeta^2} & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} \zeta_t = \hat{\zeta}_t \triangleq \begin{bmatrix} \hat{F}_{xxt}^i & \hat{F}_{x\delta t}^i \\ \hat{F}_{\delta xt}^i & \hat{F}_{\delta\delta t}^i \end{bmatrix}$$

$$\hat{\zeta}_t = E(\zeta_t/t) = \begin{bmatrix} \hat{x} \\ 0 \end{bmatrix} \quad (\text{puisque } E(\delta_t/t) = 0)$$

$$\text{On note : } \tilde{\zeta} = \zeta - \hat{\zeta}$$

$$\text{et } \hat{N}_t \triangleq E(\tilde{\zeta}_t \tilde{\zeta}_t^T/t) = \begin{bmatrix} M_t & 0 \\ 0 & \Phi_t \end{bmatrix}$$

En développant l'équation $f_t(\zeta_t)$ jusqu'au 2ème ordre autour de $\hat{\zeta}_t$, il vient :

$$\underline{x}_{t+1} \sim \underline{f}_t(\hat{\zeta}_t) + \hat{F}_{\zeta t} \tilde{\zeta}_t + \frac{1}{2} \begin{bmatrix} \tilde{\zeta}_t^T \hat{F}_{\zeta\zeta t}^1 \tilde{\zeta}_t \\ \vdots \\ \tilde{\zeta}_t^T \hat{F}_{\zeta\zeta t}^n \tilde{\zeta}_t \end{bmatrix}$$

1° Solution :

On néglige les termes du 2ème ordre.

Alors :

$$\underline{x}_{t+1} \sim \hat{F}_{xt} \underline{x}_t + \underline{f}_t(\hat{x}_t, 0) - \hat{F}_{xt} \hat{x}_t + \hat{F}_{\delta t} \delta_t$$

que nous identifions à :

$$\underline{x}_{t+1} \sim A_t \underline{x}_t + \underline{u}_t + \underline{b}_t$$

$$\text{soit } \underline{b}_t = \hat{F}_{\delta t} \delta_t$$

δ_t étant indépendant de $\underline{x}_t$, nous avons bien $\underline{b}_t$ indépendant de $\underline{x}_t$

d'où :

$$\begin{aligned} A_t &= \hat{F}_{xt} \\ \underline{u}_t &= \underline{f}_t(\hat{x}_t, 0) - \hat{F}_{xt} \hat{x}_t \\ \phi_{bt} &= \hat{F}_{\delta t} \phi_t \hat{F}_{\delta t}^T \end{aligned}$$

2° Solution :

On utilise les termes du 2ème ordre du développement de $\underline{f}_t(\underline{\zeta}_t)$

$$\text{Soit en notant } \underline{d}_t = \frac{1}{2} \begin{bmatrix} \underline{\zeta}_t^T \hat{F}_{\zeta\zeta t}^1 \underline{\zeta}_t \\ \vdots \\ \underline{\zeta}_t^T \hat{F}_{\zeta\zeta t}^n \underline{\zeta}_t \end{bmatrix}$$

$$\underline{x}_{t+1} = \underline{f}_t(\underline{\zeta}_t) \sim \underline{f}_t(\hat{\underline{\zeta}}_t) + \hat{F}_{\zeta t} [\underline{\zeta}_t - \hat{\underline{\zeta}}_t] + \underline{d}_t$$

Rappels :

- Si x_1, x_2, x_3, x_4 sont des variables gaussiennes centrées :

$$\Rightarrow E(x_1 x_2 x_3) = 0$$

$$\Rightarrow E(x_1 x_2 x_3 x_4) = E(x_1 x_2) E(x_3 x_4)$$

$$+ E(x_1 x_3) E(x_2 x_4)$$

2-1-13

$$+ E(x_1 x_4) E(x_2 x_3)$$

(pour une démonstration, voir par exemple Laning et Battin p. 158-159)

- Trace (AB) = Trace (BA)

(A : matrice $p \times r$, B : matrice $r \times p$)

Faisons l'hypothèse que $(\underline{\zeta}_t/t)$ est une variable gaussienne.

Une application fastidieuse des rappels ci-dessus permet de

calculer :

$$-\hat{\underline{d}}_t = E(\underline{d}_t/t) = \frac{1}{2} \begin{bmatrix} \text{Tr}(\hat{F}_{\zeta\zeta t}^1 \hat{N}_t) \\ \vdots \\ \text{Tr}(\hat{F}_{\zeta\zeta t}^n \hat{N}_t) \end{bmatrix}$$

D'autre part, notons $\hat{\underline{d}} \triangleq \underline{d} - \hat{\underline{d}}$

On trouve encore :

$$-\hat{P}_t = E(\hat{\underline{d}}_t \hat{\underline{d}}_t^T) = \frac{1}{2} \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \text{Tr}(\hat{F}_{\zeta\zeta t}^i \hat{N}_t) & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \text{Tr}(\hat{F}_{\zeta\zeta t}^j \hat{N}_t) & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

$$-0 = E(\hat{\underline{d}}_t \underline{\zeta}_t^T)$$

La prise en considération des termes du second ordre se traduit donc seulement par les corrections $\hat{\underline{d}}_t$ et $\hat{\underline{P}}_t$ sur le vecteur $\underline{u}_t$ et la matrice Φ_{bt} calculés précédemment, soit :

$$\begin{aligned} \underline{A}_t &= \hat{\underline{F}}_{xt} \\ \underline{u}_t &= \underline{f}_t(\hat{\underline{x}}_t, \underline{0}) - \hat{\underline{F}}_{xt} \hat{\underline{x}}_t + \hat{\underline{d}}_t \\ \Phi_{bt} &= \hat{\underline{F}}_{\delta t} \Phi_t \hat{\underline{F}}_{\delta t}^T + \hat{\underline{P}}_t \end{aligned} \quad 2-1-14$$

2-1-5 Discussion

Compte-tenu de (2-1-12) ou (2-1-14), les équations (2-1-8 à 2-1-11) constituent un algorithme de Filtrage Non Linéaire Approximativement optimal.

Nous appellerons F.N.L.A. cette méthode, en précisant au besoin l'ordre (premier ou second) des termes retenus pour l'établissement de la linéarisation (2-1-7).

Nous pourrions établir l'algorithme de filtrage dans le cas d'une équation de mesure non linéaire (équations 2-1-3, 2-1-4). Nous préférons garder les algorithmes sous la forme présente, plus condensée, en remarquant d'une part que les observations sont très souvent directement une combinaison linéaire d'états, d'autre part que nous proposons par la suite des modalités d'application, par sous-programmes spécialisés élémentaires, et qu'il est alors préférable d'effectuer au préalable la mise en forme du problème exposée dans le paragraphe 2-1-2.

Pour établir l'algorithme de filtrage, nous avons fait l'hypothèse que $(\underline{z}_t/t)$ est une variable gaussienne (rappelons que le vecteur $\underline{z}$ est constitué de l'état $\underline{x}$ et du vecteur $\underline{\delta}$)

En ce qui concerne $\underline{\delta}$, cette hypothèse peut être une donnée du problème.

Par contre, si $(\underline{z}_t/t)$ est gaussien, $\underline{x}_{t+1} = \underline{f}_t(\underline{z}_t)$ n'est pas gaussien (si $\underline{f}_t$ n'est pas linéaire) et donc $(\underline{z}_{t+1}/t+1)$ ne sera pas gaussien, ce qui infirme l'hypothèse et la réduit à une approximation dont nous donnerons cependant une justification au paragraphe 2-3. Cette approximation est d'autant meilleure que la linéarisation de $\underline{f}_t(\underline{z}_t)$ est mieux justifiée, c'est-à-dire que les termes du second ordre ($\underline{d}_t$) sont plus petits.

Or, $\hat{d}_t$ et $\hat{p}_t$ fournissent une mesure de la valeur standard de $(\underline{d}_t)$. Si elle tend vers zéro, on peut conclure alors que l'algorithme est *asymptotiquement optimal*. Il n'en reste pas moins que la convergence de l'algorithme n'est jamais assurée, sauf dans le cas de $\underline{f}_t$ linéaire, où la convergence est alors celle du filtre de Kalman classique. Une étude détaillée de la convergence des divers algorithmes de filtrage non linéaire approché n'est pas encore apparue et, en fait, il n'est pas clair qu'un résultat généralement utile pourrait sortir d'une telle étude (Pearson, 1968). C'est la raison pour laquelle nous insisterons sur les résultats expérimentaux pour justifier l'intérêt de la méthode étudiée.

2-2 MODALITES GENERALES D'APPLICATION DU FILTRAGE NON LINEAIRE A L'IDENTIFICATION -

2-2-1 Elaboration d'un vecteur d'état généralisé

Nous précisons ici les notions introduites dans le premier chapitre. Soit un système discret représenté par des équations d'état :

$$\begin{array}{l} \underline{y}_{t+1} = \underline{k}_t(\underline{y}_t, \underline{\beta}_t) \\ \underline{s}_{t+1} = \underline{h}_{t+1}(\underline{y}_{t+1}, \underline{\gamma}_{t+1}) \end{array} \quad \left| \begin{array}{l} \text{Processus :} \\ \text{état } \underline{y}_t, \text{ sortie } \underline{s}_t \\ \text{paramètres } \underline{\theta} \end{array} \right.$$

$\underline{\beta}_t, \underline{\gamma}_t$: variables aléatoires indépendantes centrées de covariance Φ_β et Φ_γ

$\underline{s}_t$: sortie observable

- l'indice t des fonctions $\underline{k}_t$ et $\underline{h}_t$ exprime à la fois la dépendance éventuelle de $\underline{y}_{t+1}$ et $\underline{s}_{t+1}$ d'une entrée variable connue, et la non stationnarité éventuelle du processus.

A partir des observations de la sortie $\underline{s}_t$, on désire identifier le processus, c'est-à-dire trouver la valeur des paramètres $\underline{\theta}_t$ qui définissent les fonctions $\underline{f}_t$ et $\underline{k}_t$ dont la forme est supposée connue.

Exemple :

$$y_{t+1} = \frac{\theta_t^1 y_t}{1 + \theta_t^2 (y_t)^2} + \theta_t^3 + \beta_t$$

$$s_{t+1} = \theta_{t+1}^4 y_{t+1} + \gamma_{t+1}$$

En considérant $\underline{\theta}_t = [\theta_t^1 \ \theta_t^2 \ \theta_t^3 \ \theta_t^4]^T$ comme une variable,

On pourra écrire :

$$y_{t+1} = k'_t(\underline{\theta}_t, y_t, \beta_t)$$

$$s_{t+1} = h'_{t+1}(\underline{\theta}_{t+1}, y_{t+1}, \gamma_{t+1})$$

Le système est stationnaire si :

$$\underline{\theta}_{t+1} = \underline{\theta}_t$$

Cette équation peut être considérée comme une équation d'état pour constituer le système :

$$\begin{cases} y_{t+1} = k'_t(\underline{\theta}_t, y_t, \beta_t) \\ \underline{\theta}_{t+1} = \underline{\theta}_t \end{cases}$$

$$s_{t+1} = h'_{t+1}(\underline{\theta}_{t+1}, y_{t+1}, \gamma_{t+1})$$

$$\text{Soit en notant } \begin{bmatrix} y_t \\ \underline{\theta}_t \end{bmatrix} = \underline{x}_t :$$

$$\underline{x}_{t+1} = \underline{f}_t(\underline{x}_t, \beta_t)$$

$$s_{t+1} = \underline{g}_{t+1}(\underline{x}_{t+1}, \gamma_{t+1})$$

(équations de la forme 2-1-3, 2-1-4)

Le problème de l'identification de $\underline{\theta}$, par l'intermédiaire de $\underline{x}$, est donc posé en termes de filtrage.

Dans le cas où les paramètres $\underline{\theta}$ évoluent dans le temps (système non stationnaire), il est nécessaire de définir la loi d'évolution de $\underline{\theta}$. En vue d'utiliser les techniques de filtrage non linéaire, il est souhaitable que cette loi d'évolution puisse être mise sous forme d'un système d'équations d'état récurrentes stochastiques :

$$\underline{z}_{t+1} = \underline{\rho}_t(\underline{z}_t, \underline{\alpha}_t)$$

$$\underline{\theta}_{t+1} = \underline{\mu}_{t+1}(\underline{z}_{t+1}, \underline{\alpha}_{t+1})$$

($\underline{\alpha}_t$: variables indépendantes centrées de covariance connue)

Nous admettons qu'on ne se propose pas d'identifier la loi d'évolution de $\underline{\theta}$, c'est-à-dire que les fonctions $\underline{\rho}_t$ et $\underline{\mu}_t$ sont connues, ainsi

que la matrice $\Phi_{\alpha t}$ des variances - covariances de α_t . Il est immédiat qu'on peut encore considérer un système d'équations :

$$\underline{x}_{t+1} = \underline{f}_t(\underline{x}_t, \underline{\beta}_t, \underline{\alpha}_t)$$

$$\underline{s}_{t+1} = \underline{g}_{t+1}(\underline{x}_{t+1}, \underline{y}_{t+1}, \underline{\alpha}_{t+1})$$

pour lesquelles l'estimation de l'état $\underline{x}_t$ apporte la solution du problème de l'identification de $\underline{\theta}$.

Dans le cas où l'on dispose de peu d'informations sur la loi d'évolution de $\underline{\theta}$, on se contentera le plus souvent d'écrire :

$$\underline{\theta}_{t+1} = \underline{\theta}_t + \underline{\alpha}_t$$

Les variations aléatoires α_t^i des composantes θ_t^i seront généralement supposées indépendantes les unes des autres (matrice Φ_{α} diagonale), et leurs variances choisies à priori en fonction de la rapidité de variation supposée des paramètres.

2-2-2 Initialisation de l'algorithme de filtrage

Pour mettre en oeuvre le filtrage non linéaire approché, il est nécessaire d'initialiser les équations récurrentes (2-8 à 2-11) par des valeurs $\hat{\underline{x}}_0$ et $\hat{M}_0$.

Par définition, $\hat{\underline{x}}_0$ est l'espérance de l'état conditionnellement aux mesures jusqu'à l'instant 0, et $\hat{M}_0$ la matrice de covariance correspondante.

En l'absence de mesures antérieures, $\hat{\underline{x}}_0$ et $\hat{M}_0$ seront considérés comme une "information à priori" sur l'état généralisé à l'instant 0, c'est-à-dire sur l'état proprement dit du processus et sur les paramètres à identifier.

Dans l'hypothèse gaussienne, on peut relier immédiatement $\hat{\underline{x}}_0$ et $\hat{M}_0$ à un intervalle de confiance sur les grandeurs concernées.

Soit par exemple $[b_i, B_i]$ l'intervalle de confiance à 95,5 % (connaissance à priori) sur une composante x^i

$$\text{On écrira : } \hat{x}_0^i = \frac{b_i + B_i}{2}, \quad \hat{M}_0^{ii} = \left(\frac{B_i - b_i}{4} \right)^2$$

Deux points de vue s'opposent alors :

1°) On dispose réellement de connaissances à priori significatives. Il est alors souhaitable de les utiliser, et la situation sera d'autant meilleure que les intervalles de confiances à priori seront plus petits.

2°) On désire tenir compte seulement des mesures. Il est clair que pour obtenir ce résultat, on doit supposer un intervalle de confiance à priori de largeur infinie, c'est-à-dire prendre comme initialisation :

$$\hat{\underline{x}}_0 \text{ quelconque } (= \underline{0} \text{ par exemple})$$

$$\hat{\underline{M}}_0 = k \begin{bmatrix} 1 \end{bmatrix} \text{ avec } k = \infty$$

Dans le cas du filtrage linéaire, on est limité dans le choix de k seulement par des questions matérielles (valeur maximale acceptable par l'organe de calcul).

Lors du filtrage non linéaire approché, nous avons fait l'hypothèse que la fonction $\underline{x}_1 = f_1(\underline{x}_0, \delta_0)$ pouvait être approchée par un développement limité de $\underline{f}_1$.

D'autre part, l'hypothèse que les estimations conditionnelles successives sont gaussiennes n'est pratiquement vérifiée que si l'approximation linéaire de l'équation d'état est acceptable dans le domaine de vraisemblance de $\underline{x}$. Si ce domaine est trop étendu, les hypothèses ci-dessus seront fortement contredites et l'algorithme, non seulement sera loin d'être optimal, mais pourra même diverger. Il est donc nécessaire de trouver un compromis entre une initialisation aussi large que possible (pour réduire l'influence des connaissances à priori) et une initialisation suffisamment étroite pour garantir le bon fonctionnement de l'algorithme.

L'expérience montre cependant que la liberté de manoeuvre est encore très étendue entre les deux extrêmes où l'algorithme diverge à la suite d'une initialisation trop large et où il reste marqué par une initialisation trop étroite.

Le problème de l'initialisation se pose d'ailleurs différemment selon qu'on envisage une estimation en ligne ou une estimation hors ligne.

Lors d'une estimation en ligne, on dispose généralement de connaissances à priori assez précises sur le système dont on désire suivre l'évolution des paramètres. D'autre part, si les mesures se poursuivent pendant un temps assez long, l'initialisation se traduit par des transitoires dont l'influence disparaît au bout de quelque temps.

Lors d'une estimation hors ligne, l'influence des connaissances à priori disparaît également si la dimension de l'échantillon est assez grande, mais de toute façon, il est possible alors de répéter plusieurs fois l'identification en recherchant une initialisation convenable, c'est-à-dire telle que l'algorithme converge, mais aussi telle que la matrice de covariance finale

sur les paramètres stationnaires à identifier, soit très inférieure à la matrice de covariance initiale. On doit également vérifier que l'estimation finale se trouve bien dans l'intervalle de confiance défini initialement.

A ces conditions, on peut considérer que l'estimation finale ne dépend pas de l'initialisation choisie. D'autre part, nous proposerons une correction finale (voir 3-3) qui permet d'éliminer approximativement (et même rigoureusement dans le cas du filtrage linéaire) l'influence des connaissances à priori.

2-3 DIFFERENTS POINTS DE VUE SELON LESQUELS ON PEUT CONSIDERER LE MEME ALGORITHME D'IDENTIFICATION -

Nous comparons ici l'algorithme d'identification par F.N.L.A. avec une forme itérative de la méthode du maximum de vraisemblance (ou de la méthode du modèle) et nous présentons les conclusions qu'on peut tirer de ce rapprochement.

Rappel des résultats de la méthode du maximum de vraisemblance

On considère un système stationnaire discret

$$\underline{y}_{t+1} = A\underline{y}_t + B\underline{u}_t + \underline{\beta}_t$$

$$\underline{s}_{t+1} = C\underline{y}_{t+1} + \underline{\gamma}_{t+1}$$

$\underline{\beta}$ et $\underline{\gamma}$: séquences indépendantes gaussiennes stationnaires centrées de covariance Φ_β et Φ_γ .

Soit un couple d'entrées-sorties : $\underline{u}_t, \underline{s}_t$ ($0 < t < T$), l'état initial $\underline{y}_0$ étant connu.

La méthode du Maximum de Vraisemblance consiste à déterminer :

- l'ensemble des coefficients $\underline{\theta}$ des matrices A, B, C (paramètres à identifier)
- une séquence $\underline{\beta}_t$ et une séquence $\underline{\gamma}_t$ satisfaisant les équations ci-dessus,

de sorte qu'un critère J soit minimum :

$$J = -\text{Log}(L) = \frac{1}{2} \sum_{t=1}^T \left[\underline{\beta}_{t-1}^T \Phi_\beta^{-1} \underline{\beta}_{t-1} + \underline{\gamma}_t^T \Phi_\gamma^{-1} \underline{\gamma}_t \right] + \frac{n}{2} \text{Log}(\det \Phi_\beta) (\det \Phi_\gamma) + \text{cte}$$

L est la vraisemblance des paramètres $\underline{\theta}$ associée à la réalisation $\underline{u}_t, \underline{s}_t$ (Schweppe, 1965).

Si les matrices Φ_β et Φ_γ sont inconnues, J doit être minimisé également par rapport aux coefficients de ces matrices.

On démontre alors, sous des conditions assez générales, que les valeurs $\hat{\theta}$, $\hat{\Phi}_\beta$ et $\hat{\Phi}_\gamma$ qui minimisent J sont des estimations asymptotiquement convergentes, efficaces et gaussiennes (Aström, 1965).

Restreignons nous ici au cas où Φ_β et Φ_γ sont connus.

La méthode du maximum de vraisemblance consiste alors à minimiser un critère *quadratique* où interviennent les "écarts d'équation" β_t et les "écarts de sortie" γ_t .

La méthode d'identification paramétrique par la *méthode du modèle* consiste, en dehors de tout contexte probabiliste, à minimiser un critère tel que :

$$J = \frac{1}{2} \sum_{t=1}^T (\beta_{t-1}^T P \beta_{t-1} + \gamma_t^T Q \gamma_t)$$

où P et Q sont des matrices définies positives arbitraires,

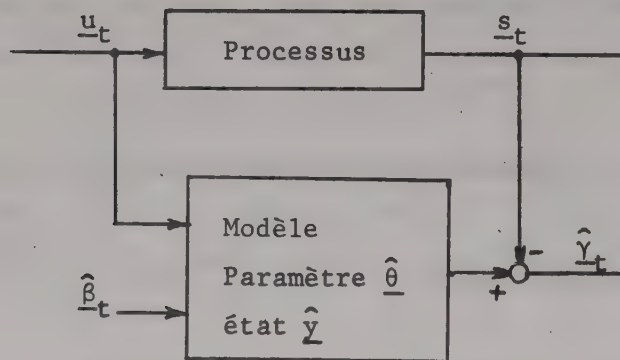


fig. (3-2)

c'est-à-dire qu'on recherche $\hat{y}$, $\hat{\theta}$, $\hat{\beta}_t$, $\hat{\gamma}_t$ obéissant aux équations :

$$\hat{y}_{t+1} = \hat{A} \hat{y}_t + \hat{B} u_t + \hat{\beta}_t$$

$$s_t = \hat{C} \hat{y}_t + \hat{\gamma}_t$$

tels que le critère J soit minimum.

Souvent d'ailleurs on ne cherche pas à tenir compte des "erreurs d'équation" β et on minimise seulement une norme (quadratique ou non)

de l'écart sortie Processus - sortie Modèle .

Il est clair que la méthode du modèle constitue une excellente approche de l'identification.

En effet, le critère adopté se suffit à lui même en dehors de toute hypothèse statistique et sa minimisation a toujours un sens. On peut même choisir un modèle dont les équations ont une forme différente de celle du processus. Ce mode d'identification rentre alors dans le cadre de la définition générale (Zadeh, 1962) :

"L'identification est la détermination, sur la base des entrées - sorties, d'un système appartenant à une classe spécifiée, auquel le système soumis à l'expérience est équivalent"

La structure du modèle (classe spécifiée) et la forme exacte de J (critère d'équivalence) peuvent être choisies arbitrairement : il est clair en effet qu'on peut considérer comme "correct" un modèle tel qu'une mesure de l'écart entre la sortie observée du processus et la sortie du modèle soit faible.

Si l'on fait l'hypothèse (d'ordre métaphysique) que la structure du modèle est identique à la structure du processus, et que les coefficients du critère correspondent à des données statistiques ($P = \Phi_{\beta}^{-1}$, $Q = \Phi_{\gamma}^{-1}$), on peut conclure que la méthode du modèle fournit une estimation statistiquement optimale (au moins asymptotiquement).

Cependant, cette hypothèse n'est pas indispensable et si les matrices Φ_{β} et Φ_{γ} adoptées pour élaborer l'algorithme du maximum de vraisemblance ne sont pas exactes, l'identification aura quand même un sens : celui de la minimisation d'un critère déterminé.

Cette double interprétation du même algorithme permet d'en renforcer la portée et, pour cette raison, il est intéressant de constater que l'identification par filtrage non linéaire peut être considérée comme une forme itérative approchée de l'identification par le maximum de vraisemblance :

Considérons en effet la séquence $\hat{\underline{x}}_t$ correspondant à l'identification récurrente de l'état généralisé d'un processus (état proprement dit $\underline{y}$ et paramètres $\underline{\theta}$).

On extrait de $\hat{\underline{x}}_t$ la séquence $\hat{\theta}_t$, $\hat{\theta}_{t+1}$...

On démontre alors (Chen, 1970) que :

$$\text{si } \hat{\theta}_t \text{ minimise } J_t = \frac{1}{2} \sum_{k=1}^t \left[\underline{\beta}_{k-1}^T \Phi_{\beta}^{-1} \underline{\beta}_{k-1} + \underline{\gamma}_k^T \Phi_{\gamma}^{-1} \underline{\gamma}_k \right]$$

$$\text{alors } \hat{\theta}_{t+1} \text{ minimise } J_{t+1} = \frac{1}{2} \sum_{k=1}^{t+1} \left[\underline{\beta}_{k-1}^T \Phi_{\beta}^{-1} \underline{\beta}_{k-1} + \underline{\gamma}_k^T \Phi_{\gamma}^{-1} \underline{\gamma}_k \right]$$

Ceci rigoureusement dans le cas du filtrage linéaire, approximativement pour le filtrage non linéaire (et alors $M_{\theta\theta} \sim (J_{\theta\theta})^{-1}$).

Notons que l'estimation par la méthode du maximum de vraisemblance étant asymptotiquement gaussienne, nous apportons une justification à l'hypothèse que les estimations conditionnelles successives sont gaussiennes, hypothèse faite au paragraphe 2-1 lors de la prise en considération des termes du 2ème ordre dans le développement des fonctions non linéaires.

Le fait de pouvoir considérer l'algorithme de filtrage non linéaire comme une formulation récurrente approchée de la méthode du maximum de vraisemblance (ou de la méthode du modèle) a une importance assez grande pour la mise en oeuvre de l'algorithme, comme pour l'interprétation des résultats.

En effet, si l'on considère le point de vue probabiliste, les variances covariances Φ_{β} et Φ_{γ} constituent une donnée du problème. Si les variances réelles ne correspondent pas aux caractéristiques présumées, on ne peut plus prétendre à l'optimalité *statistique* de l'algorithme d'estimation, mais on peut cependant considérer celui-ci comme optimal *au sens du critère J*. Inversement dans le cadre de l'identification par la méthode du modèle, il faut se fixer le critère à minimiser, c'est-à-dire chiffrer le poids relatif des différents écarts. Le point de vue statistique permet alors de donner une valeur aux coefficients P et Q du critère, en fonction des variances présumées de $\underline{\beta}$ et $\underline{\gamma}$.

Dans tous les cas, nous voyons que se pose le problème de l'estimation des variances des bruits et perturbations ($\underline{\beta}$ et $\underline{\gamma}$). Ce problème sera étudié en détail dans le chapitre 4. Répétons cependant que, même si les variances Φ_{β} et Φ_{γ} utilisées dans l'algorithme ne correspondent pas aux valeurs exactes, le point de vue déterministe permet de conclure que les coefficients identifiés sont au moins ceux qui minimisent les corrections $\underline{\beta}$ et $\underline{\gamma}$ nécessaires pour faire coïncider le comportement du modèle et celui du processus (fig. 3-2).

Chapitre 3

MODALITES PRATIQUES DE LA MISE EN OEUVRE DU FILTRAGE NON LINEAIRE

3-1 Décomposition de l'équation d'état généralisée en une succession d'opérations élémentaires.

3-2 Identification des systèmes continus.

3-3 Application "hors ligne" et "en ligne".
Élimination des connaissances à priori.

3-4 Étude d'une simulation.

Nous montrons ici comment les équations générales du filtrage non linéaire peuvent conduire à une mise en oeuvre relativement simple, aussi bien pour les systèmes discrets que les systèmes continus.

Une simulation permet de comparer les résultats du filtrage avec ceux de la méthode du maximum de vraisemblance, en tenant compte d'une correction permettant d'éliminer l'influence des initialisations.

3-1 DECOMPOSITION DE L'EQUATION D'ETAT GENERALISEE EN UNE SUCCESSION D'OPERATIONS ELEMENTAIRES -

Nous introduisons le principe de la méthode sur un exemple type d'identification :

Soit une transmittance du 2ème ordre à identifier :

$$H(Z) = \frac{Y(Z)}{U(Z)} = \frac{cZ^{-1}}{1 - aZ^{-1} - bZ^{-2}}$$

(Z^{-1} : opérateur retard)

Compte tenu des perturbations d'entrée-sortie, les équations du processus s'écriront :

$$\begin{aligned} y_{t+1} &= ay_t + by_{t-1} + cu_t + \delta_t \\ s_t &= y_t + \delta'_t \end{aligned}$$

- u_t : entrée connue
- s_t : sortie mesurée
- a, b, c : coefficients stationnaires à identifier à partir des observations s_t

Introduisons l'état généralisé : $\underline{x}$

soit :

$$\begin{cases} x_t^1 = y_t \\ x_t^2 = y_{t-1} \\ x_t^3 = a \\ x_t^4 = b \\ x_t^5 = c \end{cases}$$

(Nous rappelons que les indices supérieurs sont des indices de "composantes").

Les équations d'état généralisées s'écriront alors :

$$\begin{aligned} x_{t+1}^1 &= x_t^3 x_t^1 + x_t^4 x_t^2 + u_t x_t^5 + \delta_t \quad (= y_{t+1}) \\ x_{t+1}^2 &= x_t^1 \quad (= y_t) \\ x_{t+1}^3 &= x_t^3 \quad (= a) \\ x_{t+1}^4 &= x_t^4 \quad (= b) \\ x_{t+1}^5 &= x_{t+1}^5 \quad (= c) \end{aligned}$$

Soit :

$$\underline{x}_{t+1} = \underline{f}_t(\underline{x}_t, \delta_t)$$

Nous introduisons d'autre part un état et une équation supplémentaire, dont la nécessité apparaîtra par la suite :

$$x_{t+1}^6 = x_t^1 \quad (= y_t)$$

L'équation d'observation étant :

$$s_{t+1} = [10 \dots 0] \underline{x}_{t+1} + \delta'_{t+1} \quad (= y_{t+1} + \delta'_{t+1})$$

Soit :

$$s_{t+1} = C_{t+1} \underline{x}_{t+1} + \delta'_{t+1}$$

Plutôt que d'appliquer directement les équations générales du filtrage au système ci-dessus, nous allons considérer un système d'état $\underline{z}_q$, stricte-

ment équivalent au système primitif, mais tel que les équations

$$\underline{z}_{q+1} = f_q(\underline{z}_q, \underline{\beta}_q)$$

$$\underline{r}_{q+1} = C_{q+1} \underline{z}_{q+1} + \underline{\beta}'_{q+1}$$

correspondent à des opérations élémentaires pouvant faire l'objet de quelques sous programmes généraux auxquels il suffit de faire appel pour établir le filtre.

Considérons l'évolution de $\underline{z}_q$ depuis $q = k$ jusqu'à $q = k + 6$, selon un cycle organisé de sorte qu'à chaque fois que q augmente d'une unité, seule une composante z^i du vecteur $\underline{z}$ change, le cycle étant initialisé par :

$$\underline{z}_k = \underline{x}_k$$

$$z^6_{k+1} = z^1_k (= y_t)$$

(autres composantes inchangées)

- pas d'observation

$$z^1_{k+2} = z^3_{k+1} z^1_{k+1} (= ay_t)$$

- pas d'observation

$$z^1_{k+3} = z^1_{k+2} + z^4_{k+2} z^2_{k+2} (= ay_t + by_{t-1})$$

- pas d'observation

$$z^1_{k+4} = z^1_{k+3} + uz^5_{k+3} (= ay_t + by_{t-1} + cu_t, \text{ si } u = u_t)$$

- pas d'observation

$$z^1_{k+5} = z^1_{k+4} + \beta_{k+4} (= y_{t+1}, \text{ si } \beta_{k+4} = \delta_t)$$

- pas d'observation

$$z^2_{k+6} = z^6_{k+5} (= y_t)$$

+ observation : $r_{k+6} = z^1_{k+6} + \beta'_{k+6}$ (avec $r_{k+6} = s_{t+1}$)
et $\beta'_{k+6} = \delta'_{t+1}$

Il est clair que :

$$\underline{z}_{k+6} = \begin{bmatrix} y_{t+1} \\ y_t \\ a \\ b \\ c \\ y_t \end{bmatrix} = \underline{x}_{t+1}$$

De même, en répétant le même cycle de récurrences, on obtient :

$$\underline{x}_{t+2} = \underline{z}_{k+12} , \dots , \underline{x}_{t+j} = \underline{z}_{k+6j} , \dots$$

La décomposition adoptée ci-dessus n'est pas unique. Notons cependant que x_{t+1}^1 est fonction de x_t^2 et que x_{t+1}^2 est fonction de x_t^1 .

Quelle que soit la succession d'équations $\underline{z}_{q+1} = f_q(\underline{z}_q, \beta_q)$ (si l'on s'impose qu'une seule composante de $\underline{z}$ se trouve modifiée), nous sommes donc contraints d'introduire la composante supplémentaire z^6 , comme une "mémoire" permettant de préserver une grandeur intermédiaire.

Nous pouvons conclure qu'un filtre appliqué au système :

$$\underline{z}_{q+1} = f_q(\underline{z}_q, \beta_q) , \underline{r}_{q+1} = C_{q+1} \underline{z}_{q+1} + \beta'_{q+1} \dots$$

permettra de calculer des estimations :

$$\hat{\underline{x}}_{t+1} = \hat{\underline{z}}_{k+6} , \hat{\underline{x}}_{t+2} = \hat{\underline{z}}_{k+12} , \dots$$

Mise en oeuvre du filtre

a) Prédiction

Les équations $\underline{z}_{q+1} = f_q(\underline{z}_q, \beta_q)$ que nous avons considérées font partie d'un nombre restreint de type d'opération :

$$I \quad \boxed{z_{q+1}^i = a_q^i z_q^i + a_q^j z_q^j} \quad (\text{linéaire})$$

Autres composantes inchangées

$$II \quad \boxed{z_{q+1}^i = a_q^i z_q^i + a_q^{kl} z_q^k z_q^l} \quad (\text{non linéaire})$$

Autres composantes inchangées

Pour ces opérations, nous pouvons développer les équations de filtrage et élaborer des sous programmes ayant pour :

- Argument d'entrée : $\hat{z}_q, \hat{M}_q, a^i, i, a^j, j$ (cas I)

$\hat{z}_q, M_q, a^i, i, a^{kl}, k, l$ (cas II)

- Argument de sortie : z_{q+1}^*, M_{q+1}^*

III Dans le cas où $\boxed{z_{q+1} = z_q + v_q + \beta_q}$, v_q étant un vecteur déterministe et β_q un vecteur aléatoire indépendant de variance-covariance ϕ_q , les équations de prédiction se réduisent à :

$$z_{q+1}^* = \hat{z}_q + v_q, \quad M_{q+1}^* = \hat{M}_q + \phi_q$$

b) Estimation

Lorsqu'il n'y a pas d'observation, nous avons simplement :

$$\hat{z}_{q+1} = z_{q+1}^*$$

$$\hat{M}_{q+1} = M_{q+1}^*$$

Nous remarquons que dans l'exemple considéré, r_q est scalaire et est fonction de seulement *une* composante du vecteur d'état.

Fréquemment, r_q est une fonction de *deux* états (lorsque dans la mesure intervient une composante continue ou un bruit "coloré").

Nous considérons donc le cas d'une observation fonction de deux états seulement, soit une équation du type suivant :

$$\text{IV } \boxed{r_q = a^i z_q^i + a^j z_q^j + \beta'_q} \quad (\beta'_q \text{ de variance } \phi_q)$$

Le sous programme correspondant aura pour :

- Argument d'entrée : $z_q^*, M_q^*, r_q, a^i, i, a^j, j, \phi'_q$

- Argument de sortie : $\hat{z}_q, \hat{M}_q$ (on peut également calculer l'espérance de la variance conditionnelle de β'_q , ce dont on verra l'utilité pour l'estimation des variances).

On peut évidemment programmer les équations d'estimation correspondant au cas général d'une estimation vectorielle fonction de l'ensemble des états :

$$r_q = C_q z_q + \beta'_q$$

Le résultat partiel précédent a cependant l'avantage de limiter le volume de calcul dans son domaine d'application, où les perturbations sur le vecteur observation sont mutuellement indépendantes et où chaque composante de l'observation est fonction linéaire de deux états au plus, domaine qui s'avère très étendu dans la pratique.

Généralisation

Considérons le développement limité du second ordre d'une équation d'état :

$$\underline{x}_{t+1} = \underline{f}_t(\underline{x}_t, \underline{\delta}_t)$$

$$\text{Reprenons la notation } \underline{\zeta} = \begin{bmatrix} \underline{x} \\ \underline{\delta} \end{bmatrix} \quad \begin{array}{l} (\underline{x} : \text{dimension } n) \\ (\underline{\zeta} : \text{dimension } m) \end{array}$$

Nous avons écrit ce développement (section 2-1-4) :

$$\underline{x}_{t+1} = \underline{f}_t(\underline{\zeta}_t) + \hat{F}_{\zeta_t} \underline{\zeta}_t + \frac{1}{2} \begin{bmatrix} \underline{\zeta}_t^T & \hat{F}_{\zeta_t}^1 & \underline{\zeta}_t \\ \vdots & \vdots & \vdots \\ \underline{\zeta}_t^T & \hat{F}_{\zeta_t}^n & \underline{\zeta}_t \end{bmatrix}$$

Nous pouvons toujours le mettre sous la forme :

$$\underline{x}_{t+1}^i = \hat{a}_t^i + \sum_{j=1}^{n+m} \hat{b}_{t,j}^i \zeta_t^j + \sum_{k=1}^{n+n} \sum_{\ell=1}^{n+m} \hat{c}_{t,k,\ell}^i \zeta_t^k \zeta_t^\ell \quad 3-1-1$$

Il est donc toujours possible d'opérer une décomposition semblable à celle qu'on présente sur l'exemple précédent :

$$\begin{aligned} \underline{z}_{q+1} &= \underline{f}_q(\underline{z}_q, \underline{\beta}_q) \\ \underline{r}_{q+1} &= C_{q+1} \underline{z}_{q+1} + \underline{\beta}_q' \end{aligned} \quad (z : \text{dimension} > n)$$

où un sous vecteur de $\underline{z}_q$ est égal à $\underline{x}_0, \dots, \underline{x}_t, \underline{x}_{t+1}, \dots$ lorsque $q = 0, \dots, k, k+p, \dots$, et où les équations $\underline{f}_q$ correspondent à des opérations élémentaires pour lesquelles sont programmées une fois pour toutes les opérations de filtrage.

Nous donnons en annexe les sous programmes correspondant à des équations des types I, II, III, IV définis plus haut, ainsi que l'organigramme du programme d'estimation correspondant à l'exemple proposé.

On peut bien entendu élaborer des sous programmes globaux, par exemple correspondant à l'équation (3-1-1), mais il faut noter que dans la pratique, un grand nombre des coefficients a^i , b^{ij} ou c^{ikl} sont le plus souvent nuls. Un programme trop général impliquerait alors beaucoup de calculs inutiles.

Toutes les applications que nous présentons ont été effectuées à l'aide des sous programmes mentionnés ci-dessus et qui sont décrits dans l'Annexe II.

Nous insistons sur l'aspect purement pratique de l'application des équations de filtrage. Cet aspect nous paraît en effet, en un sens, fondamental : dans la mesure où il n'existe pas d'algorithme fini statistiquement optimal, le critère de facilité de mise en oeuvre prend une grande importance, à condition bien entendu que l'efficacité de l'algorithme ait été contrôlée par des résultats expérimentaux.

3-2 SYSTEMES CONTINUS -

Lors de l'identification par filtrage non linéaire, les paramètres s'ajoutent à l'état proprement dit pour constituer un état généralisé dont la dimension peut donc atteindre rapidement une dimension importante.

D'autre part, compte tenu de la matrice de variance-covariance, la dimension du filtre croît comme le carré de cette dimension, ce qui écarte matériellement la possibilité d'une solution par voie analogique dans le cas de l'identification de systèmes continus.

C'est principalement pour cette raison que nous ne développons pas ici les équations d'un filtre continu, équations qui sont plus utiles dans le cadre du filtrage proprement dit lorsque la dimension de l'état ne dépasse pas quelques unités.

Nous nous proposons alors de déduire des équations d'état continues un système discret auquel nous pourrions appliquer le filtre précédemment décrit.

Parmi les différents problèmes qui peuvent se rencontrer, nous retenons un cas qui se présente fréquemment :

- Système continu représenté par une équation différentielle stochastique :

$$dx = g(x,t) dt + dw$$

- Observations discrètes :

$$s(t_k) = C_k x(t_k) + \delta_k$$

- $\underline{w}(t)$: processus de Wiener indépendant

$$E(d\underline{w} d\underline{w}^T) = P dt$$

- $\underline{\delta}'_k$: séquence aléatoire indépendante centrée

$$E(\underline{\delta}'_k \underline{\delta}'_k^T) = \Phi'$$

- Période d'échantillonnage : $t_{k+1} - t_k = T$, supposée petite vis-à-vis de la rapidité d'évolution du système.

Nous noterons : $\underline{x}_k = \underline{x}(t_k)$

$$\underline{G}_x = \frac{\partial \underline{g}}{\partial \underline{x}} (\underline{x}_k, t_k)$$

$$\underline{g}_t = \frac{\partial \underline{g}}{\partial t} (\underline{x}_k, t_k)$$

Pour appliquer les méthodes précédentes, on se propose de remplacer l'équation différentielle par une équation discrète : $\underline{x}_{k+1} = \underline{f}_k(\underline{x}_k) + \underline{\delta}_k$

a) Approximation du 1er ordre

$$\underline{x}_{k+1} = \underline{x}_k + \int_{t_k}^{t_k+T} \underline{g}(\underline{x}, t) dt + \int_{t_k}^{t_k+T} d\underline{w} \quad 3-2-1$$

Soit l'approximation :

$$\boxed{\underline{x}_{k+1} = \underline{x}_k + T \underline{g}(\underline{x}_k, t_k) + \underline{\varepsilon}_k} \quad 3-2-2$$

où $\underline{\varepsilon}_k$ est une séquence aléatoire indépendante centrée de variance-

covariance $\boxed{\Phi_{\varepsilon} = PT}$

b) Approximation du 2ème ordre

Dans l'équation (3-2-1), on remplace $\underline{g}(\underline{x}, t)$ par le développement :

$$\underline{g}(\underline{x}, t) \sim \underline{g}(\underline{x}_k, t_k) + \underline{G}_x [\underline{x} - \underline{x}_k] + \underline{g}_t [t - t_k]$$

Le facteur $[\underline{x} - \underline{x}_k]$ peut lui-même être approché par :

$$\underline{x} - \underline{x}_k \sim \underline{g}(\underline{x}_k, t_k) [t - t_k] + \int_{t_k}^t d\underline{w}$$

d'où

$$\underline{x}_{k+1} = \underline{x}_k + \int_{t_k}^{t_k+T} \left[\underline{g}(\underline{x}_k, t_k) + \{ G_x \underline{g} + \underline{g}_t \} (t - t_k) \right] dt$$

$$+ \int_{t_k}^{t_k+T} G_x \int_{t_k}^t \underline{dw} dt + \int_{t_k}^{t_k+T} \underline{dw}$$

Soit en notant $\underline{y}_k = G_x \int_{t_k}^{t_k+T} \int_{t_k}^t \underline{dw} dt + \int_{t_k}^{t_k+T} \underline{dw}$

$$\underline{x}_{k+1} \sim \underline{x}_k + T \underline{g} + \frac{T^2}{2} (G_x \underline{g} + \underline{g}_t) + \underline{y}_k$$

3-2-3

Les trois premiers termes du second membre forment l'expression classique d'intégration numérique d'une équation d'état $\dot{\underline{x}} = \underline{g}(\underline{x}, t)$ déterministe, par approximation au 2ème ordre (Cf. Boudarel-Delmas-Guichet, 1969).

Il faut encore évaluer la matrice $\Phi_{\gamma k}$ des variance-covariance de $\underline{y}_k$

On peut écrire : $\underline{y}_k = \underline{\gamma} (t_k + T)$

avec
$$d\underline{y} = \begin{cases} d\underline{y}_1 = \underline{dw} \\ d\underline{y}_2 = \underline{y}_1 dt \end{cases} = \begin{bmatrix} 0 & : & 0 \\ \dots & & \dots \\ 1 & : & 0 \end{bmatrix} \underline{y} dt + \begin{bmatrix} 1 \\ \dots \\ 0 \end{bmatrix} \underline{dw} = \underline{A} \underline{y} dt + \underline{B} \underline{dw}$$

$$\underline{y} = \underline{y}_1 + G_x \underline{y}_2 = \begin{bmatrix} 1 & : & G_x \end{bmatrix} \underline{y} = \underline{C} \underline{y}$$

Sachant que :

$$\underline{y}_1(t_k) = \underline{0}$$

$$\underline{y}_2(t_k) = \underline{0}$$

On a bien

$$\underline{y}_1 = \int_{t_k}^t \underline{dw}$$

$$\underline{y}_2 = \int_{t_k}^t \int_{t_k}^{\tau} \underline{dw} d\tau$$

et donc $\gamma(t_k+T) = \gamma_k$

On calcule alors :

$$E [\underline{y}(t_k+T) \underline{y}^T(t_k+T)]$$

On sait que l'équation d'évolution des variances-covariances de y pour un système $dy = Aydt + Bdw$ s'écrit :

$$\dot{M} = AM + MA^T + BPB^T$$

(Cf. Papoulis, 1965)

d'où en intégrant de t_k à t_k+T , à partir de $M(t_k) = 0$

$$M(t_k+T) = \begin{bmatrix} PT & | & \frac{PT^2}{2} \\ \hline \frac{PT^2}{2} & | & \frac{PT^3}{3} \end{bmatrix}$$

Sachant que

$$E(\underline{y} \underline{y}^T) = CMC^T$$

et en éliminant les puissances de T supérieures à deux, on obtient :

$$\boxed{\phi_{\gamma k} \sim PT + (G_x P + P G_x^T) \frac{T^2}{2}} \quad 3-2-4$$

c) Approximation d'ordre supérieur ou égal à 2

D'une façon générale, on peut écrire une approximation discrète :

$$\underline{x}_{k+1} = \underline{f}_k(\underline{x}_k) + \underline{\delta}_k \quad 3-2-5$$

Cependant l'expression de $\underline{f}_k$ et ϕ_k risque de devenir complexe dès que l'approximation choisie dépasse l'ordre 1, et donc d'alourdir considérablement les calculs du filtre spécialement celui des variances-covariances $\hat{M}$ et M^* .

Pratiquement, on peut cependant envisager certaines simplifications pour la mise en oeuvre du filtre :

Ecrivons le développement du 2ème ordre de l'équation (3-2-5) :

$$\underline{x}_{k+1} = \underline{f}_k(\underline{x}_k) + \hat{F}_{xk} \underline{\hat{x}}_k + \begin{bmatrix} \vdots \\ \underline{\hat{x}}_k^T \hat{F}_{xxk} \underline{\hat{x}}_k \\ \vdots \end{bmatrix} + \underline{\delta}_k \quad 3-2-6$$

Quelle que soit l'expression choisie pour $f_k(\underline{x}_k)$, nous proposons de prendre pour $\hat{F}_{xk}$, $\hat{F}_{xxk}^i$ et $\hat{\Phi}_k$, les valeurs déduites de l'approximation au premier ordre (3-2-2).

En d'autres termes, nous pourrions écrire :

$$\underline{x}_{k+1} = \underline{x}_k + T \underline{g}(\underline{x}_k, t_k) + \underline{v}_k + \underline{\delta}_k$$

$$\text{avec } \underline{v}_k = \left[\underline{f}_k(\hat{\underline{x}}_k) - \hat{\underline{x}}_k - T \underline{g}(\hat{\underline{x}}_k, t_k) \right]$$

$$\text{et } \hat{\Phi} = TP$$

Le vecteur correctif $\underline{v}_k$, non aléatoire conditionnellement à t_k , est déterminé en calculant $\underline{f}_k(\hat{\underline{x}}_k)$ par une méthode quelconque d'intégration analytique ou numérique.

L'augmentation de complexité du filtre portera uniquement sur le calcul de $\underline{x}_{t+1}^*$ et non sur le calcul de M_{t+1}^* .

On voit que l'expression 3-2-6 est exacte lorsque $\hat{\underline{x}}_k = \underline{x}_k$, malgré les approximations sur $\hat{F}_x$ et $\hat{F}_{xx}^i$, ce qui n'occasionnera donc pas de biais dans l'algorithme d'estimation.

3-3 APPLICATION "HORS LIGNE" ET "EN LIGNE"

ELIMINATION DES CONNAISSANCES A PRIORI -

En raison de la forme récurrente de l'algorithme, l'application naturelle du filtrage non linéaire à l'identification est l'estimation "en ligne" des paramètres d'un processus, stationnaire ou non.

Dans tous les cas, le terme "en ligne" implique que toutes les observations ne sont pas mises en mémoire depuis l'instant initial jusqu'à l'instant présent, et sont remplacées par une statistique de dimension finie qui, dans le meilleur des cas est une statistique *exhaustive* : la connaissance de toutes les mesures antérieures est équivalente à la connaissance de cette statistique.

L'intérêt d'une estimation en ligne n'est pas nécessairement la possibilité d'opérer en temps réel afin d'élaborer par exemple un contrôle adaptatif du processus selon l'évolution des paramètres.

L'avantage pourra être parfois simplement d'utiliser les données au fur et à mesure de leur apparition sans qu'il soit nécessaire de les mémoriser dans leur ensemble ; et même si cette mémorisation est possible, on peut trouver

intérêt à une méthode d'exploitation permettant de traiter toutes les données en un seul passage.

Par opposition, considérons la méthode du maximum de vraisemblance : celle-ci se traduit par la minimisation d'une fonction non linéaire des mesures. Cette minimisation doit être effectuée par une méthode de programmation non linéaire qui impose le calcul réitéré de la valeur du critère (à partir de la totalité de l'enregistrement). Si ce nombre d'itération est important, ainsi que la dimension des observations, on pourra préférer traiter ce problème par le filtrage non linéaire qui lui est asymptotiquement équivalent.

Le résultat final de l'identification dépend cependant toujours de l'initialisation $\hat{\underline{\theta}}_0, \hat{M}_{\theta\theta_0}$. Pour se rapprocher de la solution du maximum de vraisemblance, nous nous proposons d'essayer d'éliminer cette influence.

On peut considérer que l'initialisation $\hat{\underline{\theta}}_0, \hat{M}_{\theta\theta_0}$ correspond à une mesure fictive :

$$\underline{s}_0 = A \underline{x}_0 + \underline{\delta}'_0$$

$$\text{où } A = \begin{bmatrix} 0 & \vdots & 1 \end{bmatrix}$$

$$\underline{s}_0 = \hat{\underline{\theta}}_0$$

$$E(\underline{\delta}'_0 \underline{\delta}'_0^T) = \hat{M}_{\theta\theta_0}$$

rappelons que $\underline{x}$ est un état généralisé : $\begin{bmatrix} \underline{y} \\ \underline{\theta} \end{bmatrix}$

L'estimation finale est alors l'approximation :

$$\hat{\underline{\theta}}_T \sim E(\underline{\theta} | \underline{s}_0, s_1, \dots, s_T)$$

Or on désirerait connaître :

$$\hat{\underline{\theta}} = E(\underline{\theta} | s_1, \dots, s_T)$$

c'est-à-dire l'estimation conditionnelle uniquement aux mesures effectives ($s_1 \dots s_T$).

Dans le cas idéal d'une estimation linéaire par rapport aux mesures, l'ordre des observations n'intervient pas dans le résultat final et on écrit la relation entre les estimations conditionnelles

$$\hat{\underline{x}} = E(\underline{x} | s_1, \dots, s_T) \text{ et}$$

$$\hat{\underline{x}}_T = E(\underline{x} | s_1, \dots, s_T, s_0) :$$

$$\hat{\underline{x}}_T = \hat{\underline{x}} + \hat{\underline{M}} \underline{A}^T (\underline{A} \hat{\underline{M}} \underline{A}^T + \hat{\underline{M}}_O)^{-1} (\underline{s}_O - \underline{A} \hat{\underline{x}})$$

$$\hat{\underline{M}}_T = \hat{\underline{M}} - \hat{\underline{M}} \underline{A}^T (\underline{A} \hat{\underline{M}} \underline{A}^T + \hat{\underline{M}}_O)^{-1} \underline{A} \hat{\underline{M}}$$

On peut résoudre cette équation par rapport aux grandeurs inconnues $\hat{\underline{x}}$ et $\hat{\underline{M}}$, et l'on trouve :

$$\hat{\underline{x}} = \hat{\underline{x}}_T + \hat{\underline{M}}_T \underline{A}^T (\underline{A} \hat{\underline{M}}_T \underline{A}^T - \hat{\underline{M}}_O)^{-1} (\underline{s}_O - \underline{A} \hat{\underline{x}}_T)$$

$$\hat{\underline{M}} = \hat{\underline{M}}_T - \hat{\underline{M}}_T \underline{A}^T (\underline{A} \hat{\underline{M}}_T \underline{A}^T - \hat{\underline{M}}_O)^{-1} \underline{A} \hat{\underline{M}}_T$$

Pour retrouver l'estimation *non conditionnelle* à une observation, étant donné l'estimation conditionnelle à cette observation, tout se passe comme s'il suffisait de reconsidérer cette observation, affectée d'une erreur de mesure de même variance, mais changée de signe.

Les équations ci-dessus permettent d'éliminer l'influence des connaissances à priori sur le résultat final, (ceci rigoureusement dans le cas du filtrage linéaire optimal).

Notons que si la matrice $\hat{\underline{M}}_{\theta\theta_0}$ est diagonale, la mesure fictive initiale peut être considérée comme une succession de mesures scalaires indépendantes, et l'élimination finale peut donc se faire en utilisant le programme spécialisé correspondant à des observations scalaires indépendantes.

Cette correction n'aura d'effet sensible que si la covariance finale $\hat{\underline{M}}_{\theta\theta_T}$ n'est pas trop inférieure à la covariance initiale $\hat{\underline{M}}_{\theta\theta_0}$, mais dans tous les cas, elle permet de rapprocher les résultats du F.N.L.A. de ceux du maximum de vraisemblance.

3-4 ETUDE D'UNE SIMULATION -

Avant de présenter - dans le Chapitre 5 - les résultats expérimentaux obtenus sur divers processus réels, nous étudions ici le comportement de l'algorithme sur la simulation d'un processus discret, ce qui nous permet de comparer les résultats de l'identification avec les valeurs exactes connues.

3-4-1 Description du processus simulé

Le processus simulé correspond à l'exemple d'introduction de ce chapitre, c'est-à-dire à la transmittance du deuxième ordre (Fig. 3-1) :

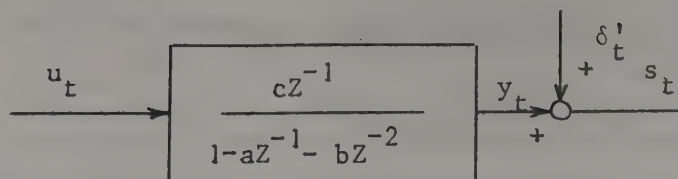


Fig. 3-1

- Valeurs numériques :

$$a = 1,4 \quad b = -0,49 \quad c = 0,09$$

- Le signal d'entrée u_t a été obtenu en filtrant une séquence indépendante gaussienne centrée δ_{ut} , de variance ϕ_u par la transmittance $1/(1-0,8Z^{-1})$. Nous obtenons ainsi le couple d'entrées sorties (u_t, s_t) représenté sur la Fig. 3-2. (On a choisi la variance ϕ_u de sorte que la variance du signal de sortie y_t soit égale à 1 en régime permanent).
- Le bruit de mesure simulé δ'_t est une séquence gaussienne centrée de variance 0,01, d'où un rapport écart type du bruit sur écart type du signal de 10 % (sauf pour le dernier essai où ce rapport sera de 100 %).
- Sauf précision contraire, les résultats sont donnés pour $T = 100$.
- Rappelons que l'organigramme du programme d'identification est donné dans l'Annexe III.

3-4-2 Sensibilité de l'algorithme aux conditions initiales

Nous considérons d'abord une initialisation "type" en prenant comme point de départ :

$$\begin{aligned} \hat{a}_0 &= 0 & \sigma_{a0} &= 10 \\ \hat{b}_0 &= 0 & \sigma_{b0} &= 10 \\ \hat{c}_0 &= 0 & \sigma_{c0} &= 10 \end{aligned}$$

Ceci correspond à des connaissances à priori relativement larges : il est légitime de supposer disponible une information initiale selon laquelle les paramètres a , b et c sont compris dans l'intervalle de confiance

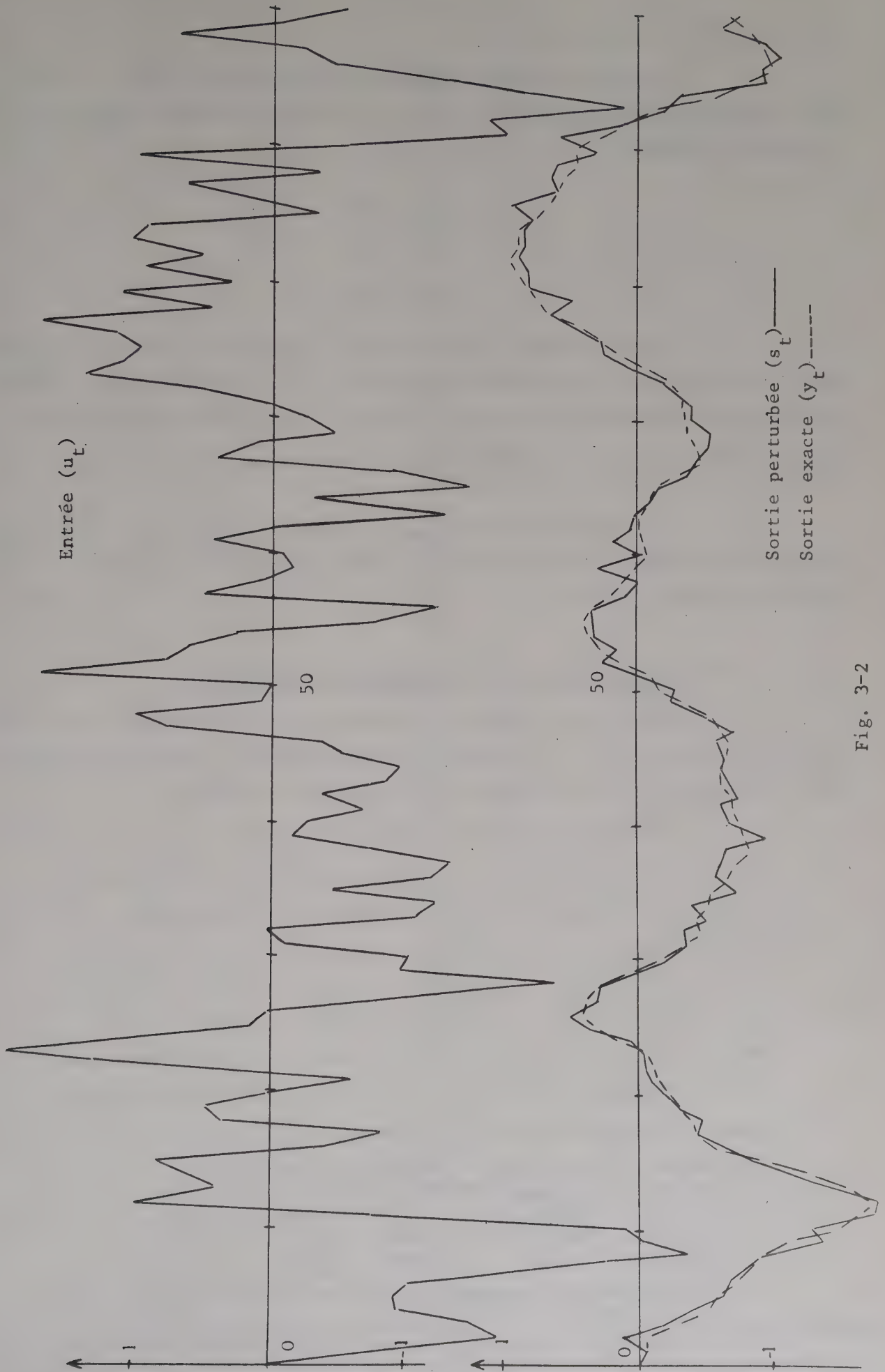


Fig. 3-2

à 95 % suivant : $[- 20, + 20]$

Les résultats sont reportés sur les Figures 3-3-a , 3-3-b et 3-3-c. Sur ces mêmes figures, on a tracé les résultats obtenus à partir de l'initialisation suivante :

$$\begin{aligned} \hat{a}_0 &= 5 & \hat{\sigma}_{a0} &= 10 \\ \hat{b}_0 &= -5 & \sigma_{b0} &= 10 \\ \hat{c}_0 &= 0,5 & \sigma_{c0} &= 10 \end{aligned}$$

Cette dernière initialisation ne correspond à rien de significatif, si ce n'est que les valeurs exactes sont bien comprises dans l'intervalle défini par l'estimation initiale ± 2 écarts type. Notons d'ailleurs que l'estimation initiale $\hat{H}_0(Z) = \frac{0,5Z^{-1}}{1-5Z^{-1} + 5Z^{-2}}$ est une transmittance instable.

On constate cependant que l'algorithme converge et qu'au bout de 30 à 40 itérations, les résultats sont pratiquement indiscernables de ceux de l'initialisation "type".

3-4-3 Comparaison avec le maximum de vraisemblance

Nous avons calculé par une méthode de programmation non linéaire les estimations $\hat{a}$, $\hat{b}$, $\hat{c}$ du maximum de vraisemblance, ainsi que l'approximation de la matrice des variances-covariances des résultats.

Les résultats du M.V. sont les suivants :

:	:	:	:	:	:	:
:	Valeurs	:	Estimation	:	Ecarts	:
:	exactes	:	:	:	type	:
:	:	:	:	:	:	:
:	a = 1,4	:	1,3808	:	0,0319	:
:	b = -0,49	:	-0,4716	:	0,0287	:
:	c = 0,9	:	0,0921	:	0,0047	:
:	:	:	:	:	:	:
:	:	:	:	:	:	:
:	:	:	:	:	:	:
:	:	:	:	:	:	:

Résultats du F.N.L.A. (à partir de l'initialisation type)

:	:	:	:	:	:	:
:	Estimation	:	écarts-type	:	Matrice de corrélation	:
:	:	:	:	:	:	:
:	1,3456	:	0,0504	:	1	:
:	-0,4377	:	0,0455	:	-0,9972	:
:	0,09720	:	0,0076	:	-0,9454	:
:	:	:	:	:	0,9334	:
:	:	:	:	:	1	:
:	:	:	:	:	:	:

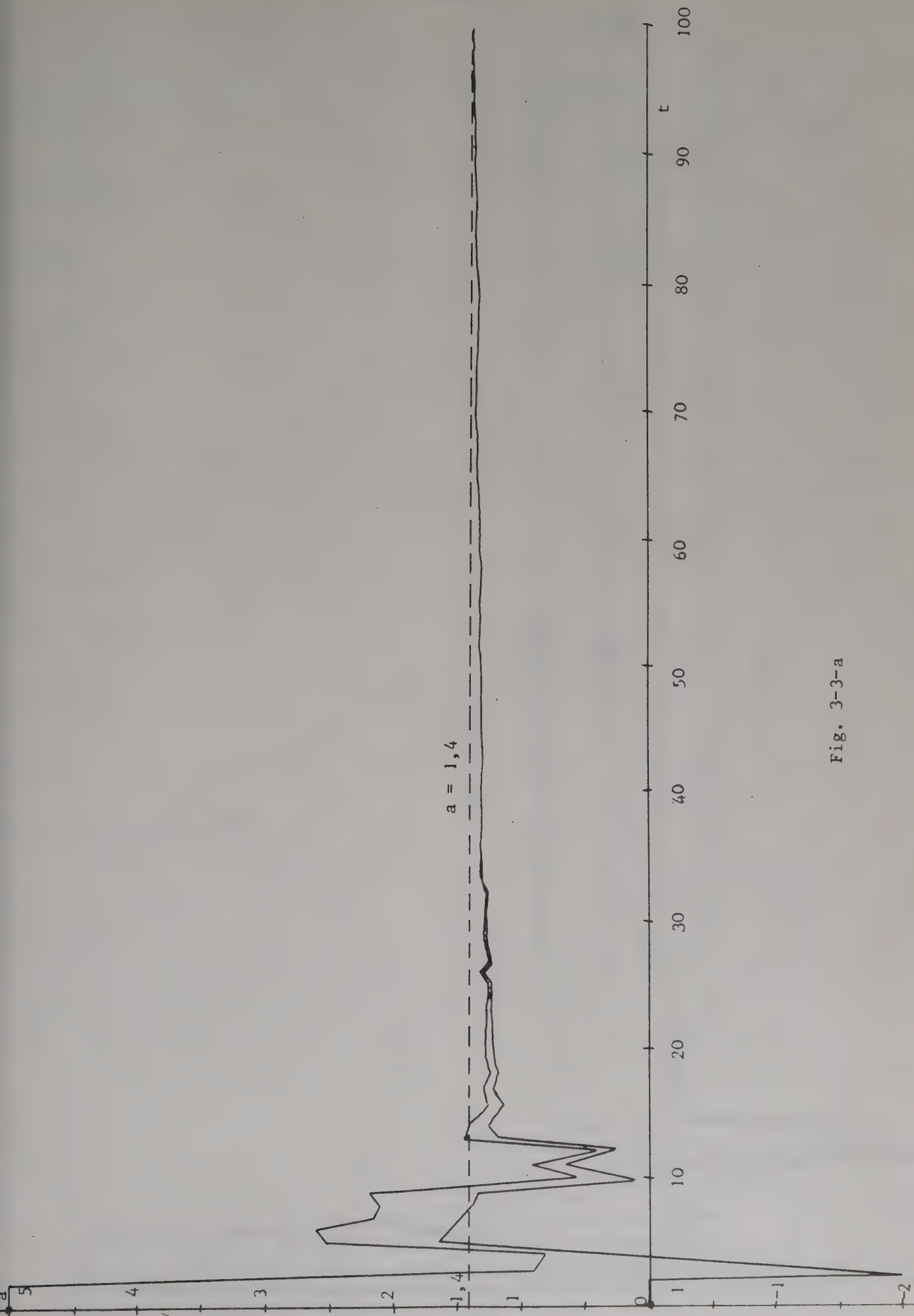


Fig. 3-3-a

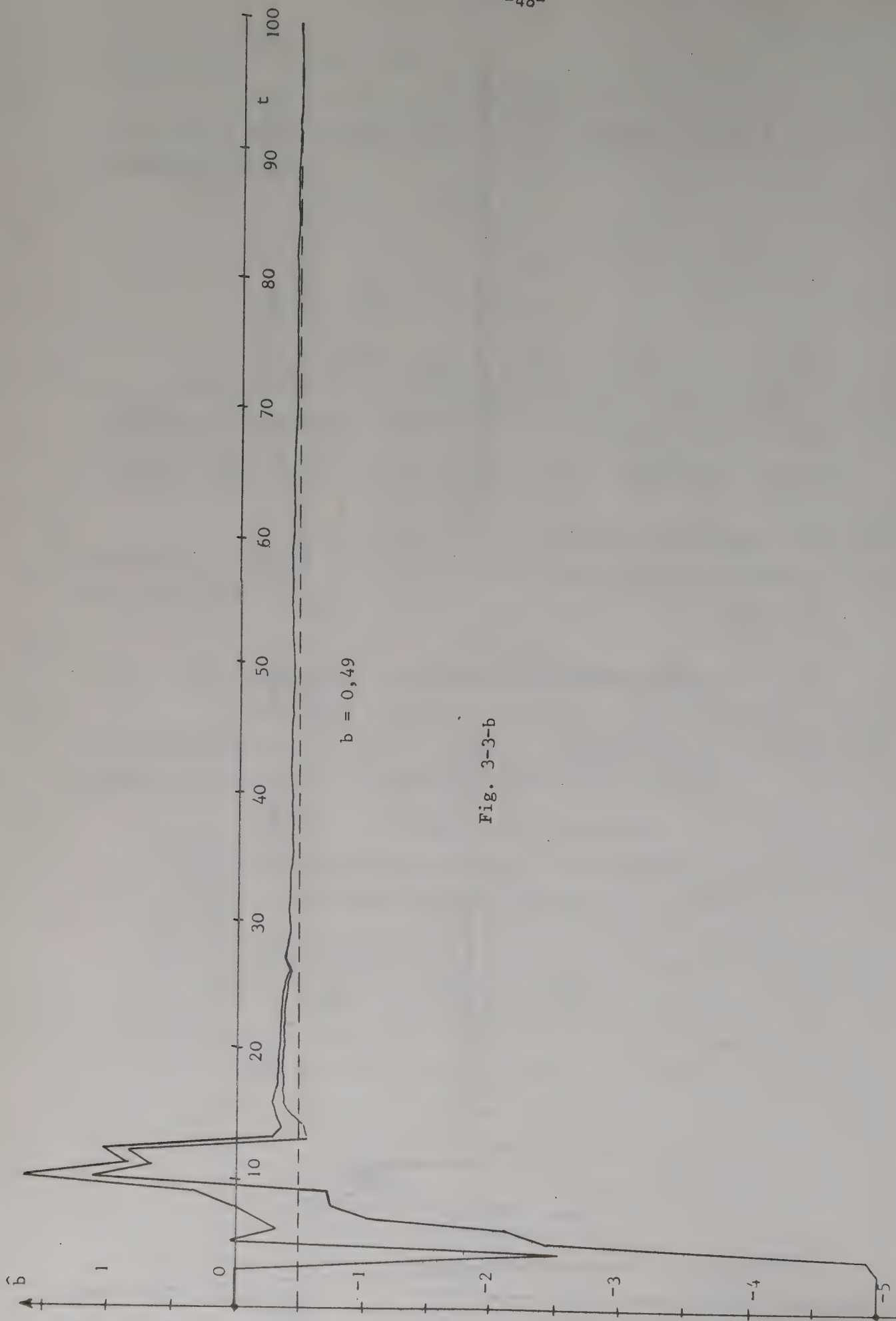


Fig. 3-3-b

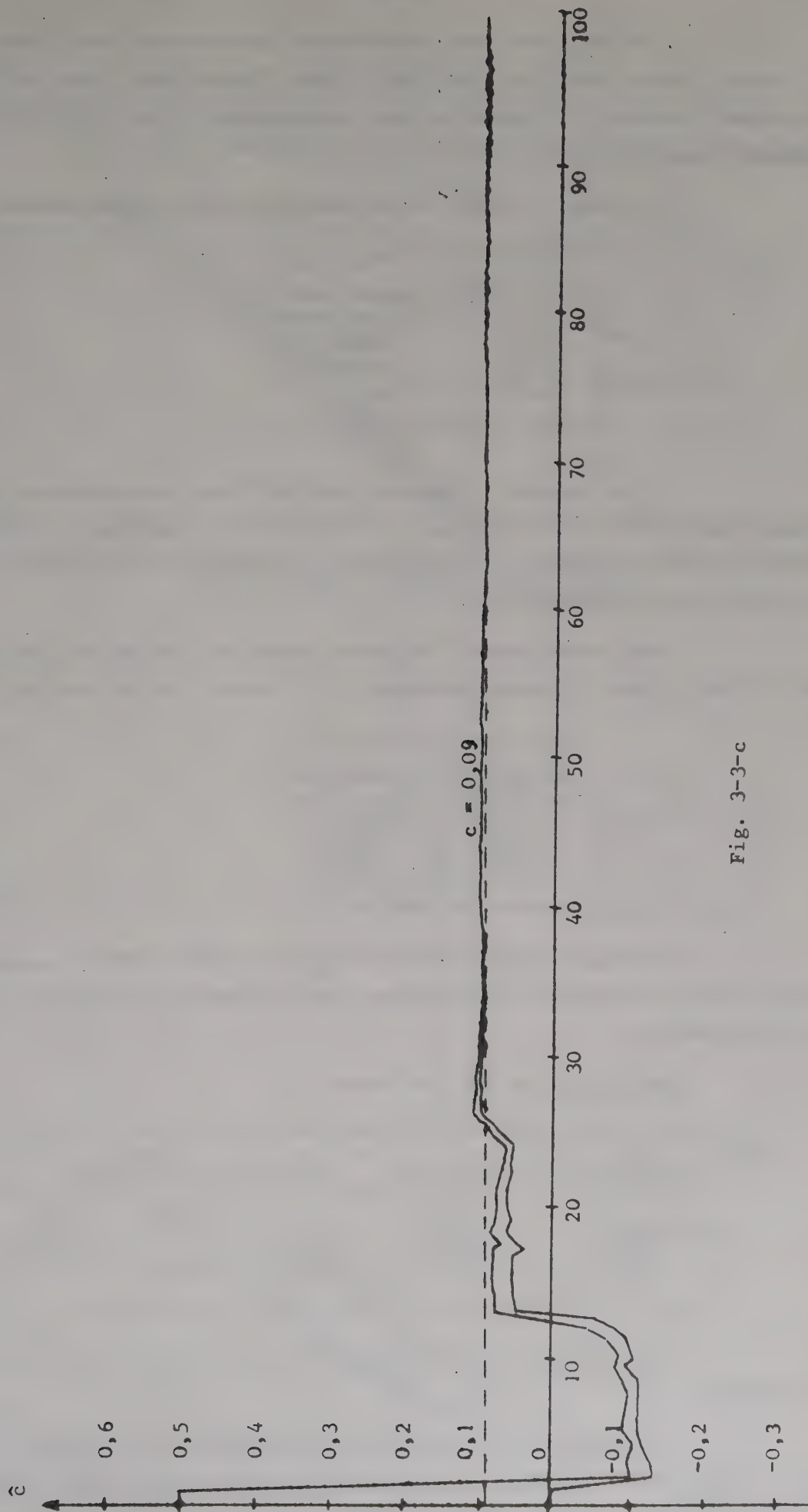


Fig. 3-3-c

Les résultats sont comparables, mais encore loin d'être identiques. Notons d'autre part, que les écarts type finals étant très inférieur aux écarts types initiaux, l'élimination des connaissances à priori n'apporte pas de changements sensibles.

On trouve en effet après élimination des connaissances à priori :

:	:	:
:	Estimation :	Ecart type :
:	-----:	-----:
:	1,3457	: 0,0504
:	-0,4377	: 0,0456
:	0,09721	: 0,0076
:	:	:

Il est possible d'obtenir une meilleure concordance en se plaçant dans des conditions "hors ligne", c'est-à-dire en reprenant l'identification avec les mêmes données, et en utilisant les résultats du précédent passage pour réinitialiser l'algorithme.

Nous choisissons directement comme initialisation les estimations et écarts types obtenus précédemment, d'où alors les estimations finales :

:	:	:
:	Estimation :	Ecart type :
:	-----:	-----:
:	1,3570	: 0,0232
:	-0,4503	: 0,0210
:	0,0953	: 0,0035
:	:	:

Les écarts types finals sont ici du même ordre de grandeur que les écarts type initiaux et l'élimination des connaissances à priori s'avère donc indispensable.

Après élimination, on trouve :

:	:	:	:	:
:	Estimation :	Ecart type :	Matrice de corrélation	
:	-----:	-----:	-----:	
:	1,3767	: 0,0363	: 1	-0,9973 -0,9396
:	-0,4679	: 0,0327	: -0,0073	1 0,9236
:	0,0924	: 0,0053	: -0,9396	0,9236 1
:	:	:	:	:

Les résultats sont cette fois très proches de ceux du maximum de vraisemblance.

A titre d'illustration, nous représentons sur la Fig. 3-4 les domaines de confiance à 95 % pour les couples de paramètres (a,b)(a,c)(b,c), à partir des résultats obtenus dans les trois cas suivants :

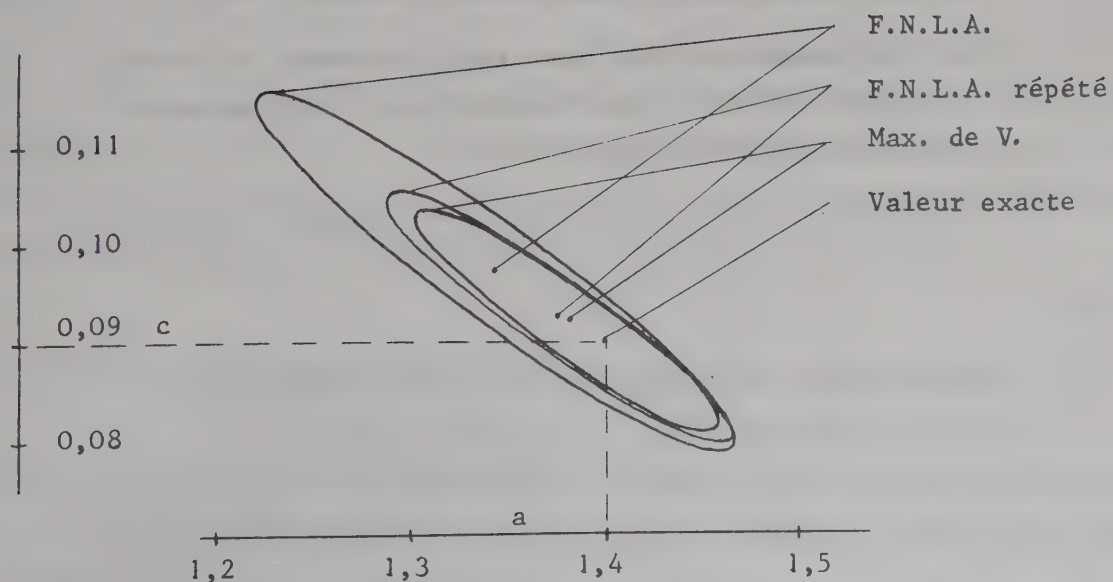
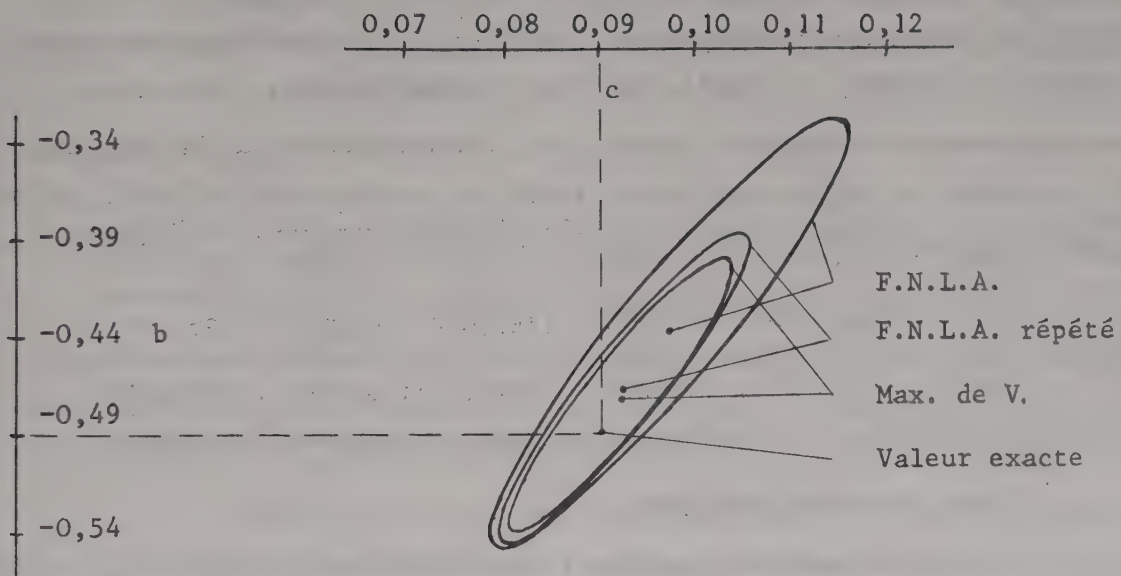
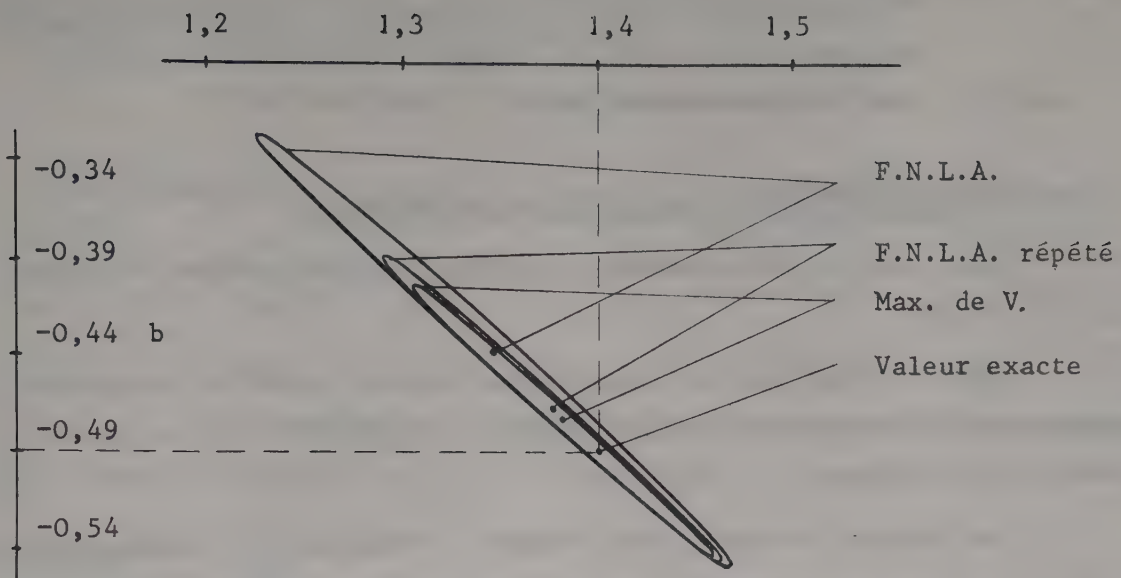


Fig. 3-4

- Maximum de vraisemblance.
- Estimation par F.N.L.A. à partir de l'initialisation type
- Estimation par F.N.L.A. répété avec élimination finale des connaissances à priori.

3-4-4 Sensibilité de l'algorithme au paramètre ϕ'

Dans tout ce qui précède, nous avons supposé connue la variance ϕ' (ici ($\phi' = 0,1$)).

Dans la réalité, il n'en est pas toujours ainsi, c'est pourquoi nous nous proposons d'observer le comportement de l'algorithme lorsqu'on affecte à ϕ' des valeurs erronées.

Le Tableau des résultats présenté ci-dessous est obtenu avec le même échantillon que précédemment, en répétant une fois l'identification avec élimination finale des connaissances à priori. On a donné successivement à ϕ les valeurs suivantes : $\phi = 0,01$, $\phi = 0,1$ (valeur exacte), et $\phi = 1$.

:	:	:	:	:	:	:	:
:	ϕ'	$\hat{a}$	$\hat{M}_{aa}$	$\hat{b}$	$\hat{M}_{bb}$	$\hat{c}$	$\hat{M}_{cc}$
:	:	:	:	:	:	:	:
:	$\phi' = 0,01$	1,3636	262.10^{-6}	-0,4569	213.10^{-6}	0,0950	625.10^{-8}
:	:	:	:	:	:	:	:
:	$\phi' = 0,1$	1,3588	146.10^{-5}	-0,4520	116.10^{-5}	0,0959	337.10^{-7}
:	:	:	:	:	:	:	:
:	$\phi' = 1$	1,3482	315.10^{-4}	-0,4416	265.10^{-4}	0,0965	698.10^{-6}
:	:	:	:	:	:	:	:

Nous observons donc que :

- les estimations proprement dites sont peu sensibles aux valeurs données à ϕ' , alors que les valeurs extrêmes de celles-ci sont dans un rapport de 1 à 100.
- les variances calculées pour les estimations ne peuvent évidemment avoir de signification que si le paramètre ϕ' a une valeur sensiblement exacte.
- comme on pouvait s'y attendre, ces variances sont approximativement proportionnelles aux valeurs affectées à ϕ' .

3-4-5 Identification en présence d'un bruit de mesure élevé

Le rapport signal/bruit est maintenant égal à 1, soit un écart type du bruit dix fois plus grand que précédemment. Toutes choses restant égales par ailleurs, on peut donc pressentir qu'il faudra bien plus

d'informations pour obtenir une précision comparable, c'est pourquoi nous considérons ici un échantillon de dimension $T = 1\ 000$.

Nous reprenons l'initialisation type et nous trouvons :

<u>Initialisation</u>		<u>Estimation finale</u>	
$\hat{a}_o = 0$	$\sigma_{ao} = 10$	$\hat{a}_T = 1,23$	$\sigma_{aT} = 0,15$
$\hat{b}_o = 0$	$\sigma_{bo} = 10$	$\hat{b}_T = 0,34$	$\sigma_{bT} = 0,13$
$\hat{c}_o = 0$	$\sigma_{co} = 10$	$\hat{c}_T = 0,116$	$\sigma_{cT} = 0,023$

Ce résultat est cohérent, les valeurs exactes étant comprises dans l'intervalle défini par $\hat{\theta}_T \pm 2$ écarts type.

CONCLUSION -

Nous avons présenté seulement quelques résultats significatifs au milieu de toutes les simulations que nous avons effectuées. De toutes façons, il ne peut être question d'entreprendre une étude expérimentale exhaustive pour un grand nombre de transmittances, de signaux d'entrée et de niveaux de bruit, d'autant plus que pour chaque caractéristique statistique de signaux, il faudrait répéter l'essai de nombreuses fois avec des réalisations différentes.

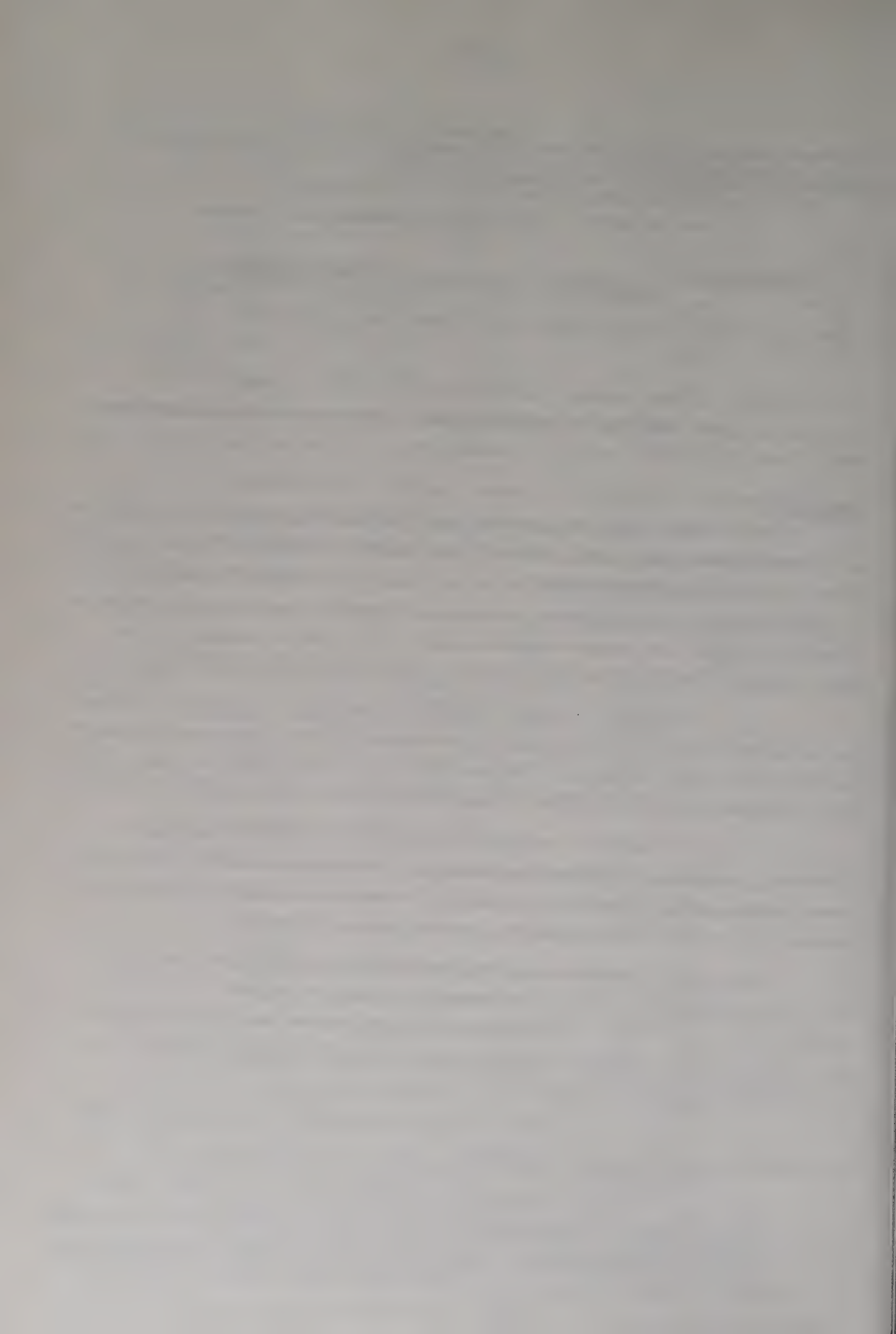
Notre série d'essais a seulement pour objet de donner un aperçu du comportement de l'algorithme dans le cas caractéristique d'une transmittance du 2ème ordre. Les conclusions qu'on peut en tirer et que, disons le bien, il n'est pas question de généraliser inconsidérément, sont les suivantes :

- Du point de vue de l'initialisation, il paraît suffisant de disposer d'un ordre de grandeur des paramètres. On peut alors initialiser systématiquement par 0, à condition de prendre un écart type initial largement supérieur à la valeur maximale estimée possible des paramètres.

- Il apparaît que, même dans sa version "en ligne", les résultats sont comparables à ceux du maximum de vraisemblance et que, lors d'une identification hors ligne, on peut améliorer l'estimation en reprenant les mêmes données avec une nouvelle initialisation, et en éliminant à posteriori les connaissances à priori introduites.

- Nous avons constaté que l'identification reste possible même en présence d'un bruit de mesure élevé.

- On observe encore qu'il n'est pas besoin de connaître exactement la variance des bruits pour obtenir une estimation correcte, mais la variance calculée pour les estimations n'aura de signification que si les variances attribuées aux bruits correspondent à la réalité, ce qui justifie l'objet du chapitre suivant consacré au problème de l'estimation des variances.



ESTIMATION DES VARIANCES

4-1 Première approche du problème.

4-2 Mise en forme telle que les caractéristiques statistiques des bruits soient identifiables par filtrage non linéaire approché.

4-3 Simplification des équations d'estimation de variance, extension au cas multivariable.

4-4 Etude d'une simulation.

Jusqu'ici, nous avons considéré comme une donnée du problème les variances Φ et Φ' des séquences indépendantes $\underline{\delta}$ et $\underline{\delta}'$ intervenant dans les équations d'état généralisées.

Bien qu'on puisse considérer, d'un autre point de vue, Φ et Φ' comme des coefficients pondérateurs d'écart dans un critère déterministe, et donc leur donner des valeurs relativement arbitraires, il est clair cependant que, plus les valeurs données à ces coefficients seront réalistes, plus le comportement de l'algorithme sera satisfaisant, et d'autre part, la matrice $\hat{M}$ n'a de sens, comme variance-covariance des estimations $\underline{x}$, que si Φ et Φ' sont bien les variances-covariances des bruits. D'où l'utilité d'une méthode permettant d'estimer ces grandeurs lorsqu'elles ne sont pas exactement connues a priori.

Nous nous proposons d'appliquer le filtrage non linéaire à ce problème, de façon à estimer simultanément $\underline{x}$ et Φ .

Nous nous limitons en effet à Φ , ayant montré précédemment (§ 2-1), qu'on pouvait toujours se placer dans le cas où $\underline{\Phi}'$ est nul.

4-1 PREMIERE APPROCHE DU PROBLEME -

Nous nous proposons d'inclure dans l'état généralisé les paramètres statistiques des bruits, afin de les identifier par filtrage non linéaire.

Il apparaît néanmoins que la mise en équation la plus naturelle conduit à un filtre non linéaire approché inefficace. C'est le point que nous

traitons dans ce paragraphe.

Reprenons les équations (2-1-5), (2-1-6) :

$$\underline{x}_{t+1} = \underline{f}_t (\underline{x}_t, \underline{\delta}_t)$$

$$\underline{s}_{t+1} = C_{t+1} \underline{x}_{t+1}$$

Nous avons admis jusqu'ici comme une donnée du problème la connaissance de la matrice de covariance :

$$\Phi_t = E(\underline{\delta}_t \underline{\delta}_t^T)$$

Nous allons considérer maintenant que les caractéristiques statistiques de la séquence $\underline{\delta}_t$ sont des paramètres à identifier.

Pour effectuer cette identification, il faut établir un nouveau système :

$$\underline{X}_t = \underline{h}_t (\underline{x}_t, \underline{\beta}_t) \quad 4-1-1$$

$$\underline{r}_t = K_t \underline{x}_t \quad 4-1-2$$

où $\underline{X}_t$ inclut l'état $\underline{x}_t$ et les paramètres statistiques de $\underline{\delta}_t$ à identifier, et où $\underline{\beta}_t$ est une séquence aléatoire indépendante dont cette fois les caractéristiques statistiques sont connues.

Pour simplifier, nous considérons ici que $\underline{\delta}_t$ est un scalaire δ_t , d'écart type σ_t à estimer.

Les résultats (négatifs) de cette première étude se généralisent immédiatement au cas vectoriel.

Nous supposons donnée l'équation d'évolution de σ_t , sous la forme :

$$\sigma_{t+1} = g_t (\sigma_t, \alpha_t)$$

où α_t est une v.a. indépendante centrée de variance $\Phi_{\alpha t}$ connue par hypothèse. Dans le cas stationnaire, cette équation se réduit à :

$$\sigma_{t+1} = \sigma_t + \alpha_t, \quad \text{où } \Phi_{\alpha t} = 0$$

Ecrivons d'autre part :

$$\delta_t = \sigma_t e_t$$

où e_t est une variable aléatoire indépendante centrée et, par définition, de variance unité.

Le système primitif (2-1-5), (2-1-6) peut alors être mis sous la forme (4-1-1), (4-1-2) :

$$\begin{cases} \underline{x}_{t+1} = \underline{f}_t (\underline{x}_t, \sigma_t, e_t) \\ \sigma_{t+1} = g_t (\sigma_t, \alpha_t) \end{cases}$$

$$\underline{s}_{t+1} = \begin{bmatrix} C_{t+1} & \vdots & 0 \end{bmatrix} \begin{bmatrix} \underline{x}_{t+1} \\ \sigma_{t+1} \end{bmatrix}$$

qui est bien de la forme désirée (4-1-1), (4-1-2) avec $\underline{x}_t = \begin{bmatrix} \underline{x}_t \\ \sigma_t \end{bmatrix}$, et $\underline{\beta}_t = \begin{bmatrix} e_t \\ \alpha_t \end{bmatrix}$.

Les variances de $\underline{\beta}_t$ sont connues, soit par hypothèse, soit par définition.

Considérons alors ce qui devient le F.N.L.A. appliqué au système $\underline{x}$ (on se référera aux notations du § 2-1-4).

La matrice A_t devient :

$$A_t = \begin{bmatrix} \hat{F}_{xt} & \hat{F}_{\sigma t} \\ \hat{G}_{xt} & \hat{G}_{\sigma t} \end{bmatrix}$$

- g n'étant pas fonction de $\underline{x}$, ses dérivées par rapport à $\underline{x}$ sont nulles, soit $\hat{G}_{xt} = 0$.

- d'autre part $F_{\sigma} = \frac{\delta f}{\delta \sigma}$ peut s'écrire $\begin{bmatrix} e & \frac{\delta f}{\partial \delta} \end{bmatrix}$

or les dérivées étant prises pour $\hat{x}$, $\hat{\sigma}$, $\hat{e}$, $\hat{\alpha}$ et que, en particulier, $\hat{e}_t = 0$, nous aurons $\hat{F}_{\sigma t} = 0$ soit finalement :

$$A_t = \begin{bmatrix} \hat{F}_{xt} & \vdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \vdots & \hat{G}_{\sigma t} \end{bmatrix}$$

Rappelons que l'observation s'écrit :

$$\underline{s}_{t+1} = \begin{bmatrix} C_{t+1} & \vdots & 0 \end{bmatrix} \begin{bmatrix} \underline{x}_{t+1} \\ \sigma_{t+1} \end{bmatrix}$$

Faisons l'hypothèse la moins restrictive : que les estimations initiales sur $\underline{x}$ et σ sont indépendantes, soit :

$$\hat{M}_0 = \begin{bmatrix} \hat{M}_{xx0} & \vdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \vdots & \hat{M}_{\sigma\sigma 0} \end{bmatrix}$$

Pour simplifier les calculs, nous nous limitons d'autre part au cas stationnaire : $\sigma_{t+1} = \sigma_t$.

On démontre alors immédiatement que :

$$\left. \begin{array}{l} \hat{\sigma}_t = \hat{\sigma}_0 \\ \hat{M}_{\sigma\sigma t} = \hat{M}_{\sigma\sigma 0} \end{array} \right\} \forall t$$

Tel que le problème est posé, on peut donc dire que σ n'est pas identifiable par le filtrage non linéaire *approché*. On doit cependant remarquer que le résultat ci-dessus est cohérent et signifie que, d'une itération à l'autre, on ne parvient pas à améliorer la connaissance de σ .

Contrairement à d'autres auteurs (Brammer, 1970) qui prennent alors en considération les moments du 3ème ordre, nous allons poser autrement le problème, de façon à le rendre soluble par le F.N.L.A.

4-2 MISE EN FORME DU PROBLEME DE SORTE QUE LES CARACTERISTIQUES STATISTIQUES DU BRUIT SOIENT IDENTIFIABLES PAR FILTRAGE NON LINEAIRE APPROCHE -

4-2-1 Nouvelle mise en équations

Nous reprenons les équations du § précédent, avec δ_t scalaire, soit :

$$\underline{x}_{t+1} = \underline{f}_t(\underline{x}_t, \delta_t)$$

4-2-1

$$\underline{s}_{t+1} = \underline{C}_{t+1} \underline{x}_{t+1}$$

δ_t : v.a. indépendante, centrée, d'écart type σ_t

L'idée directrice dans ce qui suit consiste à introduire parmi les équations d'état une nouvelle équation où intervient δ_t^2 , ce qui se justifie à priori en considérant qu'on cherche à identifier l'écart type de δ_t .

Considérons l'équation suivante, qui définit une variable aléatoire γ_t :

$$\delta_t^2 - \sigma_t^2 + \gamma_t = 0$$

γ_t est une variable aléatoire indépendante et centrée :

$$E(\gamma_t) = \sigma_t^2 - E(\delta_t^2) = 0$$

D'autre part, sa variance est :

$$E(\gamma_t^2) = E[(\sigma_t^2 - \delta_t^2)^2] = \sigma_t^4 - 2\sigma_t^2 E(\delta_t^2) + E(\delta_t^4)$$

Soit en utilisant l'égalité 2-1-12 pour calculer $E(\delta_t^4)$, on trouve :

$$E(\gamma_t^2) = 2\sigma_t^4$$

D'où l'écart type de γ_t : $\sqrt{2} \sigma_t^2$

On peut donc écrire :

$\gamma_t = \sqrt{2} \sigma_t^2 \varepsilon_t$, où ε_t est une variable aléatoire centrée, de variance unité.

L'équation ci-dessus devient alors :

$$0 = \delta_t^2 - \sigma_t^2 + \sqrt{2} \sigma_t^2 \varepsilon_t \quad 4-2-2$$

Nous allons établir un système équivalent à (4-2-1) où interviendra l'équation (4-2-2) considérée comme une équation de mesure. Il nous faut définir à quel instant on est supposé prendre connaissance du fait que le premier membre "0" de (4-2-2) est nul.

δ_t étant indépendant de $\underline{s}_t$, nous aurons toujours $(\delta_t / \underline{s}_t) = 0$, et on peut pressentir que l'observation "0" à l'instant t n'aurait alors aucune utilité sur l'objectif final qui est l'estimation de σ_t . Par contre, $\underline{s}_{t+1}$ dépend de δ_t , et donc la mesure $\underline{s}_{t+1}$ pourra permettre une estimation: $E(\delta_t / t+1)$ non identiquement nulle, et qui se répercutera sur l'estimation de σ_t .

Pour être efficace, la "mesure" devra intervenir après l'instant $t+1$, donc au moins à $t+2$, et nous écrirons donc :

$$0_{t+2} = \delta_t^2 - \sigma_t^2 + \sqrt{2} \sigma_t^2 \varepsilon_t \quad 4-2-3$$

En toute rigueur, cette équation n'apporte aucune information supplémentaire, c'est-à-dire que, dans le cadre du filtrage strictement optimal, son effet serait strictement nul.

Elle permet seulement de mettre en forme le problème en sorte que σ_t soit identifiable par F.N.L.A.

Reprenons d'autre part les équations introduites au paragraphe précédent :

$$\delta_t = \sigma_t \varepsilon_t \quad 4-2-4$$

$$\sigma_{t+1} = g_t(\sigma_t, \alpha_t) \quad 4-2-5$$

Pour mettre les équations (4-2-1) à (4-2-5) sous la forme générale souhaitée :

$$\underline{\chi}_{t+1} = \underline{h}_t (\underline{\chi}_t, \underline{\beta}_t)$$

$$\underline{r}_{t+1} = \underline{K}_t \underline{\chi}_t$$

nous allons devoir introduire quelques états supplémentaires. Ainsi, avec les notations :

$$z_t = 0_t, \quad y_t = \delta_{t-1}, \quad \rho_t = \sigma_{t-1}, \quad e'_t = e'_{t+1}, \quad \epsilon'_t = \epsilon_{t-1}$$

on obtient :

$$\left. \begin{aligned} \underline{x}_{t+1} &= f_t(\underline{x}_t, \delta_t) \\ y_{t+1} &= \delta_t \\ \delta_{t+1} &= g_t(\sigma_t, \alpha_t) e'_t \\ \sigma_{t+1} &= g_t(\sigma_t, \alpha_t) \\ \rho_{t+1} &= \sigma_t \\ z_{t+1} &= y_t^2 - \rho_t^2 + \sqrt{2} \rho_t^2 \epsilon'_t \end{aligned} \right\} \underline{\chi}_{t+1} = \underline{h}_t (\underline{\chi}_t, \underline{\beta}_t)$$

$$\left. \begin{aligned} \underline{s}_{t+1} &= C_{t+1} \underline{x}_{t+1} \\ 0_{t+1} &= z_{t+1} \end{aligned} \right\} \underline{r}_{t+1} = \underline{K}_t \underline{\chi}_t$$

Dans ces équations, δ_t est maintenant considéré comme un état, et le vecteur $\underline{\beta}_t$ est constitué des variables aléatoires e'_t , ϵ'_t , α_t dont les variances sont connues, soit par hypothèse ($\Phi_{\alpha t}$ donné), soit par définition ($\Phi'_{\epsilon t} = \Phi'_{\epsilon t} = 1$).

Il est donc possible d'appliquer l'algorithme général de F.N.L.A. au système ci-dessus.

La complexité du filtre résultant est bien moindre qu'il ne semble en considérant le nombre d'états ci-dessus.

4-2-2 Calcul du filtre

L'application des formules générales du filtre est directe, mais fastidieuse. Nous retenons essentiellement qu'il apparaît dans la matrice de covariance de $\underline{\chi}$ des blocs identiquement nuls (tels que $\hat{M}_{\sigma x}$), et il en résulte qu'on peut étudier indépendamment l'effet des diverses équations que nous avons introduites, soit successivement :

- ① $\delta_t = \sigma_t e_t$
- ② $\begin{cases} \underline{x}_{t+1} = \underline{f}_t(\underline{x}_t, \delta_t) \\ \underline{s}_{t+1} = \underline{c}_{t+1} \underline{x}_{t+1} \end{cases}$
- ③ $0 = \delta_t^2 - \sigma_t^2 + \sqrt{2} \sigma_t^2 \varepsilon_t$
- ④ $\sigma_{t+1} = g_t(\sigma_t, \alpha_t)$

Nous allons préciser le rôle de chacune des équations 1 à 4 , ce qui nous permettra progressivement de construire l'organigramme de filtrage présenté sur la Fig. (4-2-1).

Nous considérons que le cycle d'estimation débute alors qu'on dispose de :

$$\hat{\sigma}_t = E(\sigma_t/t) \text{ avec } \hat{M}_{\sigma\sigma t} = E(\hat{\sigma}_t^2/t)$$

$$\hat{x}_t = E(x_t/t) \text{ avec } \hat{M}_t = E(\tilde{x}_t \tilde{x}_t^T/t)$$

$$\text{et } M_{x\sigma t} = 0$$

① Utilisation de $\delta_t = \sigma_t e_t$

Nous calculons la variance de δ_t *conditionnellement* à t

Rappels :

- e_t est une variable indépendante de variance 1

$$- \sigma_t = \hat{\sigma}_t + \tilde{\sigma}_t$$

$$- \hat{M}_{\sigma\sigma t} = E(\tilde{\sigma}_t^2/t)$$

$$\text{Alors : } E(\delta_t^2/t) = E[(\hat{\sigma}_t e_t + \tilde{\sigma}_t e_t)^2/t]$$

$$= \hat{\sigma}_t^2 E(e_t^2/t) + 2\hat{\sigma}_t E(e_t^2 \tilde{\sigma}_t/t) + E(\tilde{\sigma}_t^2 e_t e_t/t)$$

d'où, en utilisant (2-1-12) pour estimer l'espérance des produits de 3 ou 4 variables centrées :

$$E(\delta_t^2/t) = \hat{\sigma}_t^2 + \hat{M}_{\sigma\sigma t}$$

Nous noterons $\hat{\Phi}_t$ la variance de δ_t conditionnellement à t ,

soit :

$$\hat{\Phi}_t = \hat{\sigma}_t^2 + \hat{M}_{\sigma\sigma t}$$

4-2-6

(2) Utilisation de $\left| \begin{array}{l} \underline{x}_{t+1} = f_t(\underline{x}_t, \delta_t) \\ \underline{s}_{t+1} = c_{t+1} \underline{x}_{t+1} \end{array} \right.$

Nous développons maintenant un cycle d'estimation de $\underline{x}$, en utilisant les équations ordinaires du F.N.L.A. à partir de $\underline{x}_t$ et $\hat{M}_t$, et en prenant $\hat{\Phi}_t$ pour variance de δ_t .

Ce filtre conduira donc à $\hat{\underline{x}}_{t+1}$ et $\hat{M}_{t+1}$.

Voyons comment il permet de plus le calcul de :

$$\bar{\delta}_t = E(\delta_t/t+1)$$

$$\bar{M}_{\delta\delta t} = E(\delta_t - \bar{\delta}_t)^2/t+1)$$

Nous introduisons ici la nouvelle notation : $\bar{x}$, en parallèle avec $\hat{x}$ et x^*

$$\bar{x}_t = E(x_t/t+1) \quad \text{"lissage"}$$

$$\hat{x}_t = E(x_t/t) \quad \text{"estimation"}$$

$$x_t^* = E(x_t/t-1) \quad \text{"prédiction"}$$

Nous pouvons toujours introduire l'état y :

$$\left| \begin{array}{l} \underline{x}_{t+1} = f_t(\underline{x}_t, \delta_t) \\ \underline{y}_{t+1} = \delta_t \\ \underline{s}_{t+1} = c_{t+1} \underline{x}_{t+1} \end{array} \right.$$

Si nous établissons un filtre pour le système ci-dessus, d'état $\begin{bmatrix} \underline{x} \\ \underline{y} \end{bmatrix}$, nous obtiendrons donc, outre $\hat{\underline{x}}_{t+1}$ et $\hat{M}_{t+1}$, l'estimation :

$$\hat{\underline{y}}_{t+1} = E(\underline{y}_{t+1}/t+1) = E(\delta_t/t+1)$$

soit $\bar{\delta}_t$ et $\bar{M}_{\delta\delta t}$

- ① $\delta_t = \sigma_t e_t$
- ② $\begin{cases} \underline{x}_{t+1} = \underline{f}_t(\underline{x}_t, \delta_t) \\ \underline{s}_{t+1} = \underline{c}_{t+1} \underline{x}_{t+1} \end{cases}$
- ③ $0 = \delta_t^2 - \sigma_t^2 + \sqrt{2} \sigma_t^2 e_t$
- ④ $\sigma_{t+1} = g_t(\sigma_t, \alpha_t)$

Nous allons préciser le rôle de chacune des équations 1 à 4 , ce qui nous permettra progressivement de construire l'organigramme de filtrage présenté sur la Fig. (4-2-1).

Nous considérons que le cycle d'estimation débute alors qu'on dispose de :

$$\hat{\sigma}_t = E(\sigma_t/t) \text{ avec } \hat{M}_{\sigma\sigma t} = E(\hat{\sigma}_t^2/t)$$

$$\hat{\underline{x}}_t = E(\underline{x}_t/t) \text{ avec } \hat{M}_t = E(\tilde{\underline{x}}_t \tilde{\underline{x}}_t^T/t)$$

$$\text{et } M_{x\sigma t} = 0$$

① Utilisation de $\delta_t = \sigma_t e_t$

Nous calculons la variance de δ_t *conditionnellement* à t

Rappels :

- e_t est une variable indépendante de variance 1

$$- \sigma_t = \hat{\sigma}_t + \tilde{\sigma}_t$$

$$- \hat{M}_{\sigma\sigma t} = E(\tilde{\sigma}_t^2/t)$$

$$\text{Alors : } E(\delta_t^2/t) = E[(\hat{\sigma}_t e_t + \tilde{\sigma}_t e_t)^2/t]$$

$$= \hat{\sigma}_t^2 E(e_t^2/t) + 2\hat{\sigma}_t E(e_t^2 \tilde{\sigma}_t/t) + E(\tilde{\sigma}_t^2 e_t e_t/t)$$

d'où, en utilisant (2-1-12) pour estimer l'espérance des produits de 3 ou 4 variables centrées :

$$E(\delta_t^2/t) = \hat{\sigma}_t^2 + \hat{M}_{\sigma\sigma t}$$

Nous noterons $\hat{\Phi}_t$ la variance de δ_t conditionnellement à t ,

soit :

Alors :

$$z = (\delta' + \overset{\circ}{\delta})^2 - (\sigma' + \overset{\circ}{\sigma})^2 + \sqrt{2} (\sigma' + \overset{\circ}{\sigma})^2 \varepsilon$$

On néglige le terme $\overset{\circ}{\sigma}^2 \varepsilon$, ce qui correspond à l'approximation du développement limité au 2ème ordre, d'où :

$$z = \delta'^2 + 2\delta'\overset{\circ}{\delta} + \overset{\circ}{\delta}^2 - (\sigma'^2 + 2\sigma'\overset{\circ}{\sigma} + \overset{\circ}{\sigma}^2) + \sqrt{2} \sigma'^2 \varepsilon + 2\sqrt{2} \sigma'\overset{\circ}{\sigma} \varepsilon$$

et :

$$\left| z' = \delta'^2 + M'_{\delta\delta} - \sigma'^2 - M'_{\sigma\sigma} \right.$$

- Calcul de $M'_{zz} = E(z'^2)$

$$\begin{aligned} z' &= 2\delta'\overset{\circ}{\delta} + (\overset{\circ}{\delta}^2 - M'_{\delta\delta}) - 2\sigma'\overset{\circ}{\sigma} - (\overset{\circ}{\sigma}^2 - M'_{\sigma\sigma}) \\ &\quad + \sqrt{2} \sigma' \varepsilon + 2\sigma' \sqrt{2} \overset{\circ}{\sigma} \varepsilon \end{aligned}$$

En utilisant (2-1-2) pour l'espérance du produit de 3 ou 4 variables aléatoires, il vient :

$$\left| M'_{zz} = 4\delta'^2 M'_{\delta\delta} + 2M_{\delta\delta}^{\prime 2} + 12\sigma'^2 M'_{\sigma\sigma} + 2M_{\sigma\sigma}^{\prime 2} + 2\sigma'^4 \right.$$

- Calcul de $M'_{z\sigma} = E(z' \overset{\circ}{\sigma})$

Nous trouvons :

$$\left| M'_{\sigma z} = -2\sigma' M'_{\sigma\sigma} \right.$$

Nous utilisons maintenant les résultats ci-dessus en considérant qu'avant l'information "0", nous avons :

$$\delta' = \bar{\delta}_t, \quad M'_{\delta\delta} = \bar{M}_{\delta\delta t}$$

$$\sigma' = \hat{\sigma}_t, \quad M'_{\sigma\sigma} = \hat{M}_{\sigma\sigma t}$$

D'autre part, par l'intermédiaire de $\bar{\delta}_t$, c'est la mesure s_{t+1} que l'information "0" répercute sur σ .

En conséquence, l'estimation σ'' est une estimation conditionnelle à $t+1$, soit :

$$\sigma'' = E(\sigma_{t+1}) = \bar{\sigma}_t$$

D'où les équations finales :

$$\bar{\sigma}_t = \hat{\sigma}_t + \frac{2\hat{\sigma}_t \hat{M}_{\sigma\sigma t}}{D_t} (\bar{\sigma}_t^2 + \bar{M}_{\sigma\sigma t} - \hat{\sigma}_t^2 - \hat{M}_{\delta\delta t}) \quad 4-2-7$$

$$\bar{M}_{\sigma\sigma t} = \hat{M}_{\sigma\sigma t} - \frac{4\hat{\sigma}_t^2 \hat{M}_{\sigma\sigma t}^2}{D_t} \quad 4-2-8$$

$$\text{Avec } D_t = 4\hat{\sigma}_t^2 \bar{M}_{\delta\delta t} + 2\bar{M}_{\delta\delta}^2 + 12\hat{\sigma}_t \hat{M}_{\sigma\sigma t} + 2\bar{M}_{\sigma\sigma t}^2 + 2\hat{\sigma}_t^4 \quad 4-2-9$$

④ Utilisation de $\sigma_{t+1} = g_t(\sigma_t, \alpha_t)$

A partir de $\bar{\sigma}_t = E(\sigma_t/t+1)$, l'équation ci-dessus permet de calculer $E(\sigma_{t+1}/t+1) = \hat{\sigma}_t$.

En particulier, si $\sigma_{t+1} = \sigma_t + \alpha_t$ (avec α_t centré de variance $\phi_{\alpha t}$), on aura simplement :

$$\hat{\sigma}_{t+1} = \bar{\sigma}_t \quad 4-2-10$$

$$\hat{M}_{\sigma\sigma t} = \bar{M}_{\sigma\sigma t} + \phi_{\alpha t} \quad 4-2-11$$

4-2-3 Organigrammes

L'ensemble des équations pour un cycle d'estimation est schématisé sur la Fig. (4-2-1).

Les équations reportées dans les blocs référencés ① à ④ sont les équations à partir desquelles sont établis les éléments correspondants du filtre (4-2-6 à 4-2-12).

Nous concluons donc que pour estimer l'écart type d'un bruit δ , il suffit d'adjoindre au filtre principal d'estimation de $\underline{x}$ le filtre secondaire dont nous avons donné les équations.

Le filtre principal doit être aménagé cependant de façon à fournir les grandeurs $\bar{\delta}_t$ et $\bar{M}_{\delta\delta t}$ (rappelons que $\bar{\delta}_t = E(\delta_t/s_{t+1})$). Nous avons montré que cela était toujours possible en introduisant un état supplémentaire

$$y_{t+1} = \delta_t.$$

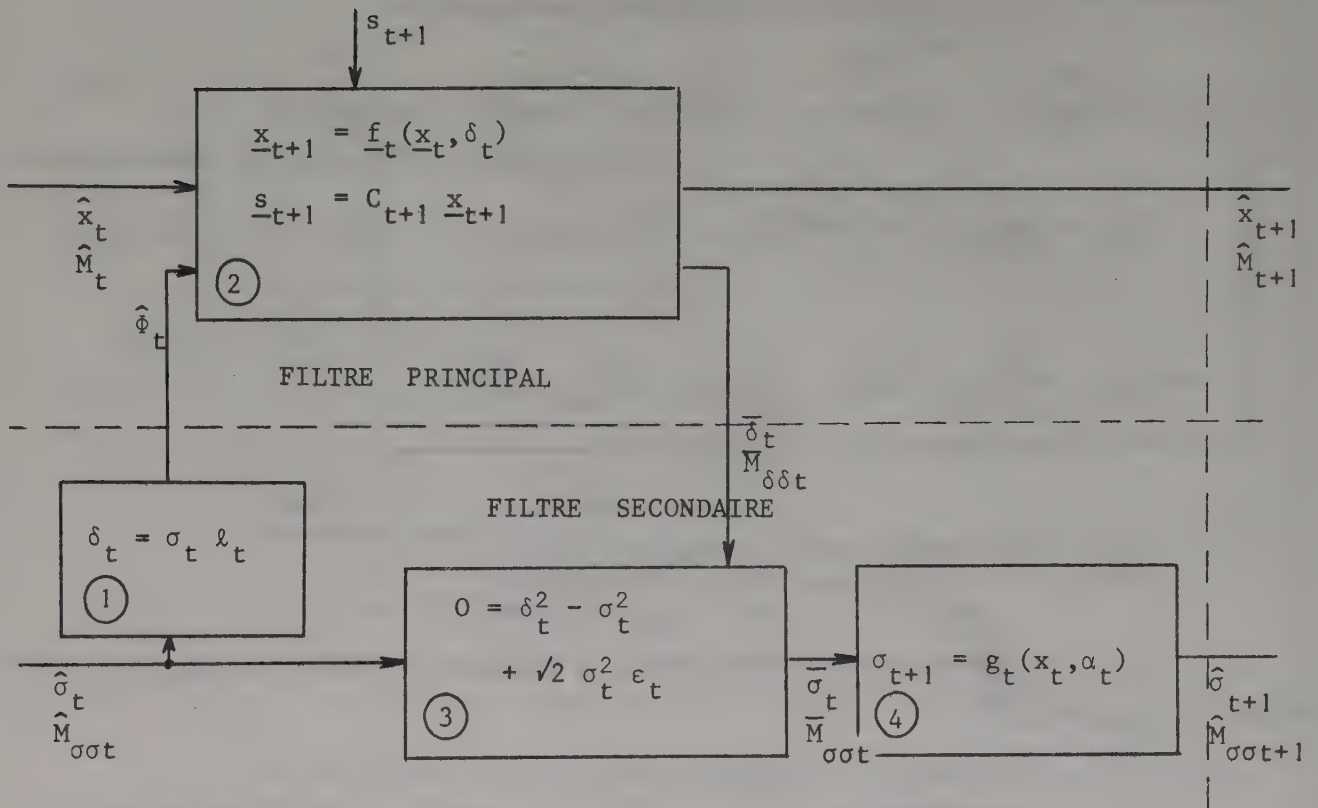


Fig. 4-2-1

Nous avons fait cette étude dans le cas d'un bruit sur l'équation d'état. Nous avons montré par ailleurs (§ 2-1-1) qu'on pouvait toujours se ramener à ce premier cas. La circonstance d'un bruit indépendant sur la mesure est cependant intéressante car nous pouvons montrer qu'il n'est pas besoin d'introduire un état supplémentaire dans le filtre principal.

Considérons par exemple :

$$\underline{x}_{t+1} = \underline{f}_t(\underline{x}_t)$$

$$\underline{s}_{t+1} = C_{t+1} \underline{x}_{t+1} + \delta_{t+1}$$

Il est clair que la mesure s_{t+1} apporte cette fois une information sur δ_{t+1} . En raison de cette distinction avec le cas d'un bruit sur l'équation d'état, l'organigramme se présentera d'une manière un peu différente (Fig. 4-2-2).

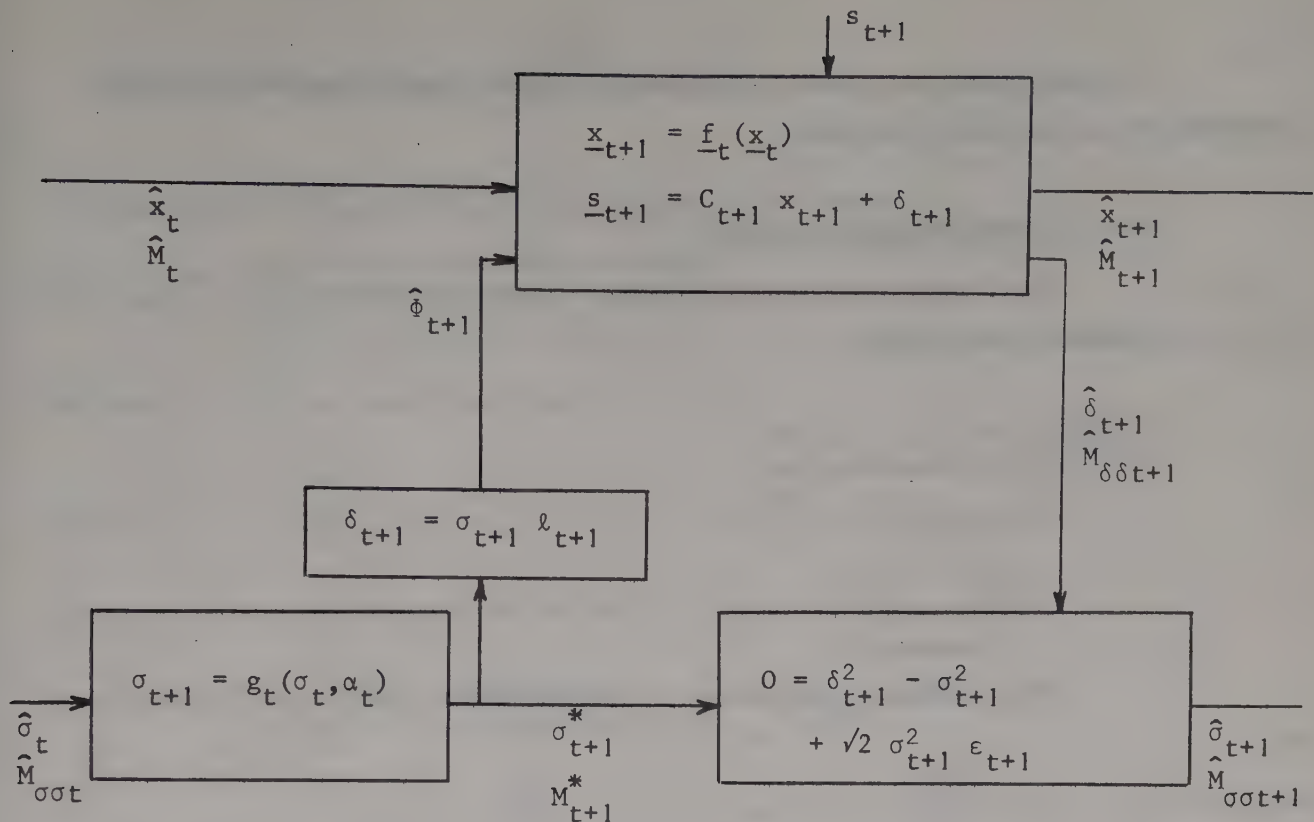


Fig. 4-2-2

Pour obtenir $\hat{\delta}_{t+1}$ et $\hat{M}_{\delta\delta t+1}$, il n'est pas besoin d'introduire un état complémentaire, et c'est pour les besoins d'une estimation de variance éventuelle que nous avons introduit dans le sous programme d'estimation le calcul de l'espérance et de la variance conditionnelle de l'erreur de mesure (voir Annexe II).

4-3 SIMPLIFICATION DES EQUATIONS D'ESTIMATION DE VARIANCE, EXTENSION AU CAS MULTIVARIABLE -

La première simplification que nous allons envisager consiste à établir les équations du filtre secondaire en considérant d'abord une linéarisation au premier ordre, en ce qui concerne σ .

En d'autres termes, on considère que l'estimation $\hat{\sigma}$ est suffisamment précise pour que dans les développements en série des fonctions non linéaires, on néglige les produits d'écart où intervient la grandeur $\tilde{\sigma} = \sigma - \hat{\sigma}$.

Cette approximation nous permettra de poser le problème directement en termes d'estimation de la variance ϕ , sans passer par l'intermédiaire de l'écart type σ .

Une deuxième approximation consistera à remplacer certains coefficients du filtre par leur comportement asymptotique.

A partir des dernières équations obtenues, nous pourrions envisager l'extension au cas multivariable.

4-3-1 Première approximation

Nous reprenons successivement les équations : (1), (2), (3),
(4) du précédent paragraphe :

$$(1) \quad \delta_t = \hat{\sigma}_t e_t + \cancel{\tilde{\sigma}_t e_t}$$

$$\text{d'où } E(\delta_t^2/t) = \hat{\sigma}_t^2$$

$$\boxed{\hat{\Phi}_t = \hat{\sigma}_t^2}$$

Nous pouvons considérer $\hat{\Phi}_t$ comme une estimation de la variance Φ_t .

$$\text{En effet : } \Phi_t = \sigma_t^2 = (\hat{\sigma}_t^2 + \tilde{\sigma}_t^2)^2 = \hat{\sigma}_t^2 + 2\hat{\sigma}_t \tilde{\sigma}_t + \cancel{\tilde{\sigma}_t^2}$$

$$\text{d'où } E(\Phi/t) = \hat{\Phi} = \hat{\sigma}_t^2$$

$$\text{et } E((\Phi_t - \hat{\Phi}_t)^2/t) = E((2\hat{\sigma}_t \tilde{\sigma}_t)^2/t)$$

$$\text{soit } \boxed{\hat{M}_{\Phi\Phi t} = 4\hat{\sigma}_t^2 \hat{M}_{\sigma\sigma t}}$$

$$(2) \quad \begin{cases} \underline{x}_{t+1} = \underline{f}_t(\underline{x}_t, \delta_t) \\ \underline{s}_{t+1} = \underline{C}_{t+1} \underline{x}_t \end{cases}$$

Nous établissons comme précédemment le filtre principal en utilisant l'estimation $\hat{\Phi}_t$

$$(3) \quad \underline{0} = \delta_t^2 - \sigma_t^2 + \sqrt{2}\sigma_t^2 \epsilon_t$$

Nous reprenons les résultats précédents en négligeant les termes $\tilde{\sigma}^2$ et $\tilde{\sigma}\epsilon$.

Les équations deviennent alors :

$$D_t = 4\hat{\delta}_t^2 \bar{M}_{\delta\delta t} + 2\bar{M}_{\delta\delta t}^2 + 2\hat{\sigma}_t^4 \quad 4-3-1$$

$$\bar{\sigma}_t = \hat{\sigma}_t + \frac{2\hat{\sigma}_t \hat{M}_{\sigma\sigma t}}{D_t} (\delta_t^2 + \bar{M}_{\delta\delta t} - \hat{\sigma}_t^2) \quad 4-3-2$$

$$\bar{M}_{\sigma\sigma t} = \hat{M}_{\sigma\sigma t} - \frac{4\hat{\sigma}_t^2 \hat{M}_{\sigma\sigma}^2}{D_t} \quad 4-3-3$$

- en remplaçant $2\hat{\sigma}_t^4$ par $2\hat{\Phi}_t^2$, l'équation (4-3-1) devient :

$$D_t = 4\bar{\delta}_t^2 \bar{M}_{\delta\delta} + 2\bar{M}_{\delta\delta}^2 + 2\hat{\Phi}_t^2 \quad 4-3-4$$

- Notons $\bar{\Phi}_t = \bar{\sigma}_t^2$ et admettons que $\bar{\sigma}_t \neq \hat{\sigma}_t$

$$\text{alors } \bar{\Phi}_t - \hat{\Phi}_t = \frac{\partial \hat{\Phi}_t}{\partial \bar{\sigma}} (\bar{\sigma}_t - \hat{\sigma}_t) = 2\bar{\sigma}_t (\bar{\sigma}_t - \hat{\sigma}_t)$$

d'où en utilisant l'équation (4-3-2)

$$\bar{\Phi}_t = \hat{\Phi}_t + \frac{4\hat{\sigma}_t^2 \hat{M}_{\sigma\sigma t}}{D_t} (\bar{\delta}_t^2 + \bar{M}_{\delta\delta t} - \hat{\sigma}_t^2)$$

ou encore, en rappelant que $\hat{M}_{\Phi\Phi t} = 4\hat{\sigma}_t^2 \hat{M}_{\sigma\sigma t}$

$$\bar{\Phi}_t = \hat{\Phi}_t + \frac{\hat{M}_{\Phi\Phi t}}{D_t} (\bar{\delta}_t^2 + \bar{M}_{\delta\delta} - \hat{\Phi}_t) \quad 4-3-5$$

- Notons de même $\bar{M}_{\Phi\Phi t} = 4\bar{\sigma}_t^2 \bar{M}_{\sigma\sigma t}$

Multiplions les deux membres de 4-3-3, le second par $4\hat{\sigma}_t^2$, le premier par

$$4\bar{\sigma}_t^2 \neq 4\hat{\sigma}_t^2$$

Il vient alors :

$$\bar{M}_{\Phi\Phi t} = \hat{M}_{\Phi\Phi t} - \frac{\hat{M}_{\Phi\Phi t}^2}{D_t} \quad 4-3-6$$

Les équations 4-3-4 à 4-3-6 constituent les nouvelles équations du filtre secondaire.

$$\textcircled{4} \quad \sigma_{t+1} = g_t(\sigma_t, \alpha_t)$$

Cette équation peut être traduite en une équation d'évolution de Φ , à moins qu'elle ne soit donnée directement sous la forme suivante :

$$\Phi_{t+1} = g'_t(\Phi_t, \beta_t)$$

Par exemple, si l'on avait :

$$\sigma_{t+1} = \sigma_t + \alpha_t$$

$$(\alpha_t \text{ de variance } \phi_{\alpha t} \ll \sigma_t^2)$$

On écrirait :

$$\sigma_{t+1}^2 = \sigma_t^2 + 2\alpha_t \sigma_t + \alpha_t^2$$

$$\text{soit : } \phi_{t+1} = \phi_t + \beta_t$$

Et l'on prendrait pour variance de β :

$$\hat{\phi}_{\beta t} = 4\sigma_t^2 \phi_{\alpha t}$$

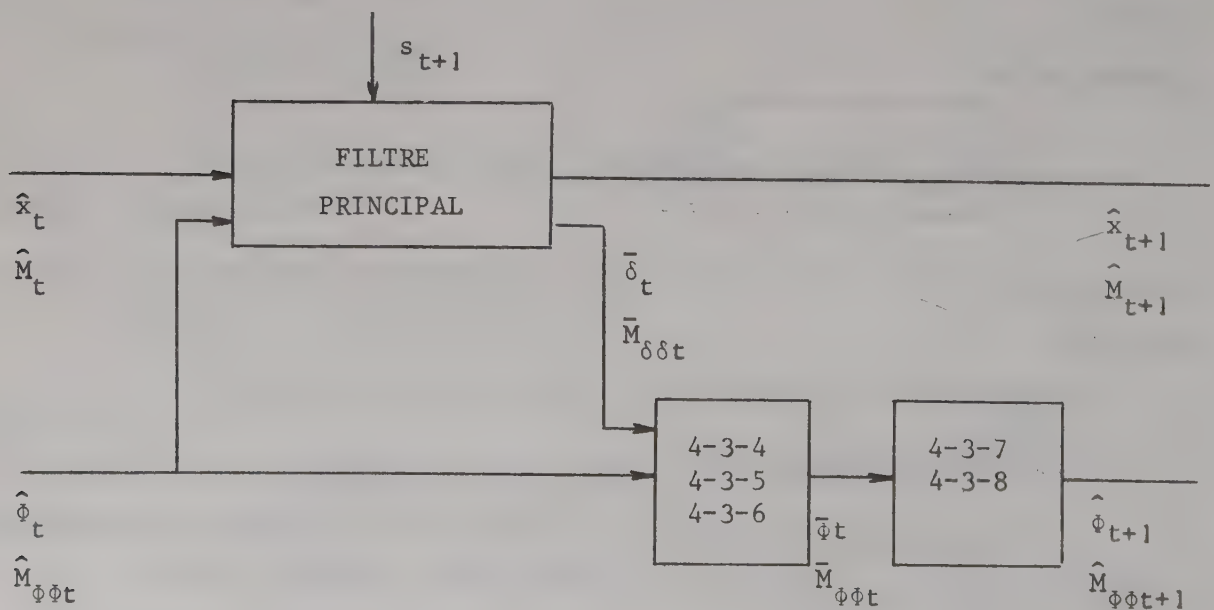
Les équations issues de (4) deviennent alors :

$$\begin{aligned} \hat{\phi}_{t+1} &= \bar{\phi}_t \\ \hat{M}_{\phi\phi t+1} &= \bar{M}_{\phi\phi t} + \hat{\phi}_{\beta t} \end{aligned}$$

4-3-7

4-3-8

Finalement, toute référence à l'écart type σ a donc disparu du filtre secondaire, et nous pouvons schématiser un cycle d'estimation :



4-3-2 Deuxième approximation

Nous proposons maintenant de remplacer les équations du filtre secondaire par leur comportement asymptotique.

Nous nous plaçons dans le cas où

$$\Phi_{t+1} = \Phi_t + \beta_t$$

Nous pouvons écrire l'ensemble des équations du filtre secondaire (4-3-4 à 4-3-8) sous la forme :

$$\begin{aligned}\hat{\Phi}_{t+1} &= \hat{\Phi}_t + \frac{\hat{M}_{\Phi\Phi t}}{D_t} (\bar{\delta}_t^2 + \bar{M}_{\delta\delta} - \hat{\Phi}_t) \\ \hat{M}_{\Phi\Phi t+1} &= \hat{M}_{\Phi\Phi t} - \frac{\hat{M}_{\Phi\Phi t}^2}{D_t} + \Phi_{\beta t}\end{aligned}$$

$$\text{Avec } D_t = 4\bar{M}_{\delta\delta}\delta_t^2 + 2\bar{M}_{\delta\delta}^2 + 2\hat{\Phi}_t^2$$

Le coefficient D_t est variable d'une itération à l'autre, mais nous pouvons considérer des valeurs extrêmes.

Remarquons en effet que $0 \leq \bar{M}_{\delta\delta} \leq \hat{\Phi}_t$ car la variance de δ_t conditionnelle à $t+1$ ne peut qu'être inférieure à la variance conditionnelle à t .

Nous trouvons alors :

- lorsque $\bar{M}_{\delta\delta t} = 0$:

$$D_t = 2\hat{\Phi}_t^2$$

- lorsque $\bar{M}_{\delta\delta t} = \hat{\Phi}_t$:

cela implique que s_{t+1} n'a apporté aucune information sur δ_t ,
et donc que $\bar{\delta}_t = \hat{\delta}_t$, or comme $\hat{\delta}_t = E(\delta_t/t) = 0$:

$$D_t = 4\hat{\Phi}_t^2$$

Nous proposons de prendre $D_t = 2\hat{\Phi}_t^2$, ce qui correspond au cas où le filtre principal fournit une "bonne" estimation de δ_t :

$$\bar{M}_{\delta\delta t} \sim 0$$

Nous obtenons alors :

$$\left| \begin{aligned} \hat{\phi}_{t+1} &= \hat{\phi}_t + \frac{\hat{M}_{\phi\phi t}}{2\hat{\phi}_t^2} (\bar{\delta}_t^2 + \bar{M}_{\delta\delta} - \hat{\phi}_t) \\ \hat{M}_{\phi\phi t+1} &= \hat{M}_{\phi\phi t} - \frac{\hat{M}_{\phi\phi}^2}{2\hat{\phi}_t^2} + \hat{\phi}_{\beta t} \end{aligned} \right|$$

Nous pouvons donner à l'algorithme une autre forme en divisant le second membre de cette dernière équation par $\hat{\phi}_t^2$, le premier par $\hat{\phi}_{t+1}^2 \neq \hat{\phi}_t^2$

$$\text{Notons : } k_t = \frac{\hat{M}_{\phi\phi t}}{\hat{\phi}_t^2}$$

$$\gamma_t^2 = \frac{\hat{\phi}_{\beta t}}{\hat{\phi}_t^2}$$

Alors :

$$\boxed{\begin{aligned} \hat{\phi}_{t+1} &= \hat{\phi}_t + \frac{k_t}{2} (\bar{\delta}_t^2 + \bar{M}_{\delta\delta} - \hat{\phi}_t) \\ k_{t+1} &= k_t - \frac{k_t^2}{2} + \gamma_t^2 \end{aligned}} \quad \begin{array}{l} 4-3-9 \\ 4-3-10 \end{array}$$

Nous retiendrons ce dernier algorithme dont la forme est particulièrement simple. Il fait apparaître des grandeurs k et γ sans dimension, dont nous allons préciser la signification et qui sont susceptibles d'exprimer les données du problème sous un point de vue favorable à leur exploitation.

- k_t représente la variance relative de l'estimation ϕ : $\left(\frac{\hat{M}_{\phi\phi t}}{\hat{\phi}_t^2} \right)$

Notons que ϕ étant une grandeur certainement positive, l'incertitude relative est toujours inférieure à 100 %, l'initialisation k_0 devra donc toujours être au plus égale à 1.

- γ_t représente la variation relative type de ϕ d'un instant à l'autre : $\left(\frac{\sqrt{\hat{\phi}_{\beta t}}}{\hat{\phi}_t} \right)$

Cette interprétation est intéressante dans la mesure où il est peu réaliste de considérer ϕ_t comme un paramètre inconnu à identifier, mais de prétendre connaître avec précision son équation d'évolution : $\phi_{t+1} = g'(\phi_t, \alpha_t)$

Sauf dans le cas stationnaire (où $\Phi_{t+1} = \Phi_t$, et donc $\Phi_{\beta t} = 0$), le paramètre γ_t , qu'on prendra en général constant : $\gamma_t = \gamma$, est celui qui nous paraît représenter le plus facilement un modèle d'évolution possible de Φ .

4-3-3 Troisième approximation

Une simplification ultime de l'algorithme consiste à remplacer k_t par sa valeur finale k .

En écrivant $k_{t+1} = k_t = k$, dans l'équation (4-3-10), on trouve :

$$k = \sqrt{2} \gamma$$

Le filtre secondaire est alors réduit à l'équation :

$$\hat{\Phi}_{t+1} = \hat{\Phi}_t + \frac{\gamma}{\sqrt{2}} (\bar{\delta}_t^2 + \bar{M}_{\delta\delta t} - \hat{\Phi}_t)$$

On peut recommander pour γ un ordre de grandeur, soit 10^{-2} , qui paraît satisfaisant pour l'identification de variances lentement variables ou même stationnaires.

4-3-4 Extension au cas multivariable

Dans le cas où δ_t n'est pas scalaire, on peut chercher une extension de la méthode précédente, consistant à identifier les coefficients d'une matrice Σ_t telle que, par définition, $\Sigma_t \Sigma_t^T = \Phi_t$.

Il apparaît encore qu'on pourrait procéder à cette identification par un filtre secondaire découplé du filtre principal, mais les équations de filtrage approché deviennent extrêmement complexes, même lorsque $\underline{\delta}$ est de dimension 2 seulement, et une formulation générale apparaît difficile à établir et ses possibilités d'application incertaines.

C'est pourquoi, nous proposons seulement de généraliser au cas multivariable les dernières expressions auxquelles nous sommes parvenus dans le cas monovarié.

Nous prendrons donc :

$$\hat{\Phi}_{t+1}^{ij} = \hat{\Phi}_t^{ij} + \frac{k_t^{ij}}{2} (\bar{\delta}_t^i \bar{\delta}_t^j + \bar{M}_{\delta\delta t}^{ij} - \hat{\Phi}_t^{ij})$$

- Si Φ est par hypothèse diagonale, on ne considérera que les équations où $i = j$.
- Dans le cas général, il faut préciser les coefficients k_t^{ij}

Si l'on dispose d'informations différentes sur les différentes composantes de $\underline{\delta}$, nous prendrons par exemple :

$$k_t^{ij} = \sqrt{k_t^i k_t^j}$$

$$\text{avec } k_{t+1}^i = k_t^i - \frac{(k_t^i)^2}{2} + \gamma^i$$

La simplification extrême consistant à prendre tous les coefficients k_t^{ij} constants et égaux à la même valeur γ .

On peut vérifier que si $\hat{\Phi}_0$ est définie positive, comme il convient à une matrice de variance-covariance, l'algorithme ci-dessus conduit à une matrice $\hat{\Phi}_t$ définie positive quel que soit t .

4-4 ETUDE D'UNE SIMULATION -

4-4-1 Position du problème

Soit un système stationnaire discret :

$$\underline{z}_{t+1} = A \underline{z}_t + \underline{\beta}_t$$

$$\underline{s}_{t+1} = C \underline{z}_{t+1} + \underline{\beta}'_{t+1}$$

4-4-1

Les coefficients stationnaires A et C étant connus, seules les variances des séquences indépendantes $\underline{\beta}_t$ et $\underline{\beta}'_t$ sont à identifier.

Considérons que le but final de cette identification est de déterminer la matrice K d'un filtre de Kalman stationnaires d'estimation de $\underline{z}$, soit :

$$\underline{z}_{t+1}^* = A \underline{z}_t^* + K (\underline{s}_t - C \underline{z}_t^*)$$

4-4-2

Plutôt que d'identifier d'abord les matrices Φ_β et Φ'_β et de calculer ensuite K par les équations de Kalman, nous nous proposons de poser directement le problème en termes d'identification de K , par l'intermédiaire de la forme "Innovation" (Kaylath, 1968). Celle-ci consiste à remarquer que la sortie $\underline{s}_t$ peut être représentée d'une manière équivalente par les équations :

$$\underline{z}_{t+1} = A \underline{z}_t + K \underline{y}_t$$

4-4-3

$$\underline{s}_{t+1} = C \underline{z}_{t+1} + \underline{y}_{t+1}$$

où K est le gain du filtre stationnaire optimal.

Sous cette forme, nous pouvons nous proposer d'identifier K et Φ_Y , au lieu de Φ_β et Φ'_β .

L'intérêt de cette représentation est qu'elle est unique. On démontre en effet (Mehra, 1970) que pour des caractéristiques statistiques données de $\underline{s}_t$, il existe plusieurs déterminations de Φ_β et Φ'_β , tandis que les coefficients de la forme innovation sont uniques.

4-2-2 Application au cas scalaire

Nous avons simulé les équations 4-4-1, où maintenant toutes les grandeurs sont scalaires, avec les valeurs numériques suivantes :

$$A = a = 0,8$$

$$\Phi_\beta = 0,6$$

$$\Phi'_\beta = 1$$

Les coefficients de la forme innovation sont donnés par les expressions :

$$\Phi_Y = \frac{1}{2} (\Phi_\beta + \Phi'_\beta (1+a)^2 + \sqrt{[\Phi_\beta + \Phi'_\beta (1+a^2)]^2 + 4\Phi_\beta'^2 a^2}) = 1,6$$

$$K = a \left(1 - \frac{\Phi'_\beta}{\Phi_Y}\right) = 0,3$$

Nous avons donc à identifier les coefficients $\Phi_Y = 1,6$ et $K = 0,3$ à partir de la séquence observée s_t (la dimension de l'échantillon est $T = 1\ 000$).

- Mise en équation

Nous modifions les équations 4-4-3 comme suit :

$$\left. \begin{aligned} \underline{z}_{t+1} &= A \underline{z}_t + K_t \underline{y}_t \\ \underline{y}_{t+1} &= \delta_t \\ K_{t+1} &= K_t \\ \underline{s}_{t+1} &= C \underline{z}_{t+1} + \underline{y}_{t+1} \end{aligned} \right\} \quad \begin{aligned} \underline{x}_{t+1} &= f(\underline{x}_t, \delta_t) \\ \underline{s}_{t+1} &= C_{t+1} \underline{x}_{t+1} \end{aligned}$$

Nous incluons ainsi le paramètre K à identifier dans l'état généralisé, et nous considérons γ_t comme un état.

Nous sommes donc ramenés à la forme 2-1-5, 1-1-6 où la variance $\Phi = \Phi_\gamma$ de δ est à identifier.

4-4-3 Résultats de l'identification

Initialisations

- Sachant que K est toujours compris entre 0 et 1, nous prendrons systématiquement :

$$\begin{aligned}\hat{K}_0 &= 0,5 && (\text{valeur exacte } 0,3) \\ \hat{\sigma}_{K0} &= 0,5\end{aligned}$$

- Nous initialiserons Φ avec des valeurs comprises entre des extrêmes 0,25 et 4 (valeur exacte 1,6)

Nous étudierons le comportement des divers filtres secondaires suivant :

- Filtre I : estimation de $\sigma = \sqrt{\Phi}$ (cf. 4-2-3)
- Filtre II : estimation de Φ par les équations simplifiées (cf. 4-3-2) :

$$\left| \begin{aligned}\hat{\Phi}_{t+1} &= \hat{\Phi}_t + \frac{k_t}{2} (\bar{\delta}_t^2 + \bar{M}_{\delta\delta} - \hat{\Phi}_t) \\ k_{t+1} &= k_t - \frac{k_t^2}{2} + \gamma_t^2 \\ &(\text{avec ici } k_0 = 0,5, \quad \gamma_t = 0)\end{aligned}\right.$$

- Filtre III : estimation de Φ avec les dernières équations simplifiées (cf. 4-3-3) :

$$\left| \begin{aligned}\hat{\Phi}_{t+1} &= \hat{\Phi}_t + \frac{\gamma}{\sqrt{2}} (\bar{\delta}_t^2 + \bar{M}_{\delta\delta} - \hat{\Phi}_t) \\ &(\text{avec } \gamma = 0,01 \text{ et } 0,001)\end{aligned}\right.$$

Les résultats sont les suivants :

- L'identification de K est, pour un échantillon $\{s_t\}$ donné, pratiquement indépendante du filtre secondaire adopté.

Sur la figure 4-4-1, nous avons représenté un tracé caractéristique d'évolution de K.

- Identification de ϕ

Avec le filtre I, nous obtenons les résultats suivants pour diverses initialisations :

$\hat{\phi}_0$	$\hat{\sigma}_{\phi 0}$	$\hat{\phi}_T$	$\hat{\sigma}_{\phi T}$
0,25	0,25	1,51	0,065
1	1	1,54	0,066
4	4	1,55	0,066

Soit une sensibilité aux conditions initiales relativement faible.

Nous comparons d'autre part, sur la Fig. 4-4-1 l'évolution de $\hat{\phi}$, à partir d'une même initialisation $\phi_0 = 1$, dans les cas suivants :

Filtre I
 Filtre II
 Filtre III $\begin{cases} \gamma = 0,01 \\ \gamma = 0,001 \end{cases}$

Nous ne constatons pas de différence notable entre I et II, par contre, pour le dernier filtre, nous observons que $\gamma = 0,01$ offre un meilleur compromis entre rapidité de convergence et incertitude finale que $\gamma = 0,001$.

Indiquons que même dans le cas où ϕ converge difficilement, K converge normalement, sans qu'on observe de différence caractéristique avec le tracé de la Fig. 4-4-1.

- Résultats obtenus avec des échantillons différents.

Nous présentons sur la Fig. 4-4-2 les résultats obtenus à partir de différents échantillons $\{s_t\}$, pour $T = 1\ 000$ et avec le filtre secondaire II.

Autant qu'on puisse en juger sur seulement 8 résultats, ceux-ci sont centrés autour de la valeur exacte, et les estimations de K sont indépendantes de celles de ϕ , ce qui justifie le découplage entre filtre principal et filtre secondaire.

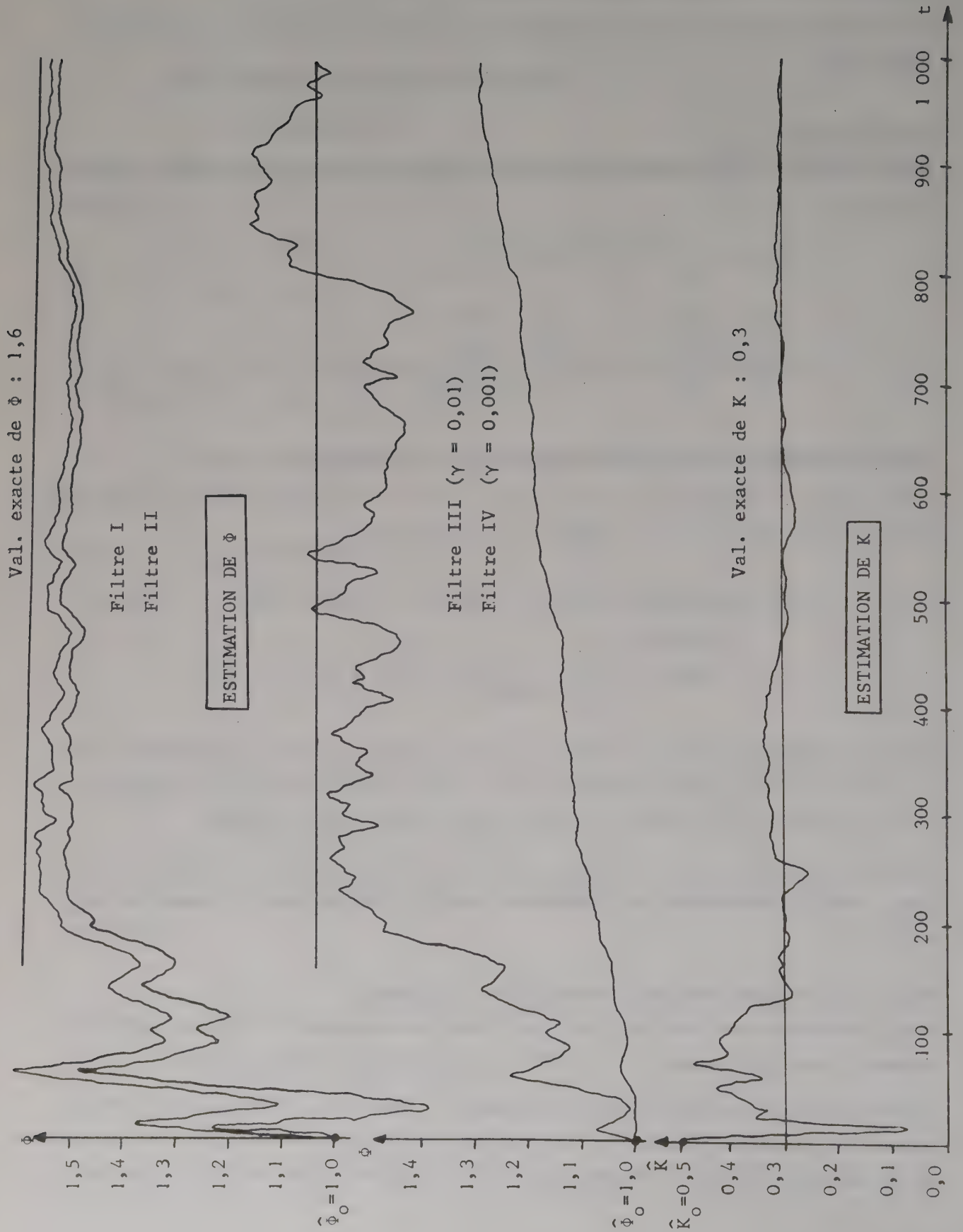


Fig. 5-4-1

K

0,5

○

Initialisation

$$\hat{K}_0 = 0,5$$

$$\hat{\sigma}_{K0} = 0,5$$

$$\hat{\Phi}_0 = 1$$

$$\hat{\sigma}_{\Phi 0} = 0,5$$

0,4

0,3

Estimations finales
pour divers échantillons

Domaine de confiance
à 95 % (centré autour de
la valeur exacte)

$$(\sigma_{KT} = 0,026, \sigma_{\Phi T} = 0,072)$$

Valeur exacte

0,2

0,1

0

1

1,6

Φ

Fig. 4-5-2



Chapitre 5

RESULTATS EXPERIMENTAUX

- 5-1 Identification d'un micro-réseau de distribution d'électricité.
- 5-2 Identification d'une batterie d'extraction liquide - liquide.
- 5-3 Identification de la chaîne de tangage d'un hélicoptère.
- 5-4 Identification d'un turbo réacteur
- 5-5 Identification d'une unité pilote générateur de vapeur.

Le premier sujet que nous traitons (5-1) apparaît comme un exemple type et permet de détailler les principaux caractères de l'identification par filtrage non linéaire : influence des connaissances à priori, identification de variance, comparaison avec l'estimation du maximum de vraisemblance. Chacun des autres exemples présente quelque difficulté particulière : bruit coloré de grande amplitude (5-2), système instable avec état initial inconnu (5-3), système non linéaire (5-4), système multivariable (5-5).

Nous ne nous étendons pas ici sur d'autres problèmes que nous avons pu traiter : en particulier, la chaîne de tangage d'un avion (type Piaggio) qui a permis de tester la discrétisation de processus continu par approximation d'ordre supérieur à 1, décrite en (3-2) ; Nous signalons également l'identification en temps réel par ordinateur en ligne de la chaîne de surchauffe du générateur de vapeur pilote du laboratoire d'Automatique de l'E.N.S.M.

Il est à noter que ces problèmes ont tous été traités en utilisant la même série de sous programmes (annexe II) qui permet d'adapter rapidement la théorie du filtrage non linéaire à chaque cas particulier.

5-1 IDENTIFICATION D'UN MICRO-RESEAU DE DISTRIBUTION D'ELECTRICITE -

(Données communiquées par le Centre de Recherches de l'E.D.F. - Clamart)

5-1-1 Présentation du problème

Il s'agit là d'un problème type, en ce sens que se trouvent

réunis les principaux éléments nécessaires pour tester les propriétés d'un algorithme d'identification, mais sans présenter de difficultés particulières.

En effet, nous considérons que :

- le système est linéaire
- décrit par une équation récurrente du 1er ordre (deux coefficients à identifier)
- avec état initial connu (conditions initiales nulles)
- bruit blanc sur la sortie, dont on peut apprécier la variance au vu de l'enregistrement.

Le couple d'entrée-sortie $\{u_t\}$, $\{s_t\}$ utilisé pour l'identification est représenté sur la figure 5-1-1.

La caractérisation adoptée est :

$$\begin{aligned} x_{t+1} &= a x_t + b u_t & x_t &: \text{état} \\ s_t &= x_t + \delta'_t & \delta'_t &: \text{bruit de mesure} \end{aligned}$$

5-1-2 Connaissances à priori

- En observant le régime permanent, on peut considérer que le gain statique $K = \frac{b}{1+a}$ est de l'ordre de -3.

- Par contre, en l'absence d'échelon sur l'entrée, la constante de temps τ du système est plus difficile à apprécier, mais on peut imaginer qu'elle est comprise entre 5 et 50 unités.

Sachant que $\tau \sim \frac{1}{1+a}$, on en déduit :

$$1^\circ) 0,8 < -a < 0,98$$

$$2^\circ) 0,06 < -b < 0,6$$

- D'autre part, l'écart type σ' du bruit de sortie est à première vue de l'ordre de 0,3. On peut admettre que 0,1 est certainement trop petit et 1 certainement trop grand.

Soit :

$$3^\circ) 0,1 < \sigma' < 1$$

5-1-3 Résultats de l'identification

a) Usage normal des connaissances à priori

Nous initialisons l'algorithme avec :

$$\left| \begin{array}{l} \hat{a}_0 = -0,9 \\ \hat{\sigma}_{ao} = 0,1 \end{array} \right| \left| \begin{array}{l} \hat{b}_0 = 0 \\ \hat{\sigma}_{bo} = 1 \end{array} \right|$$

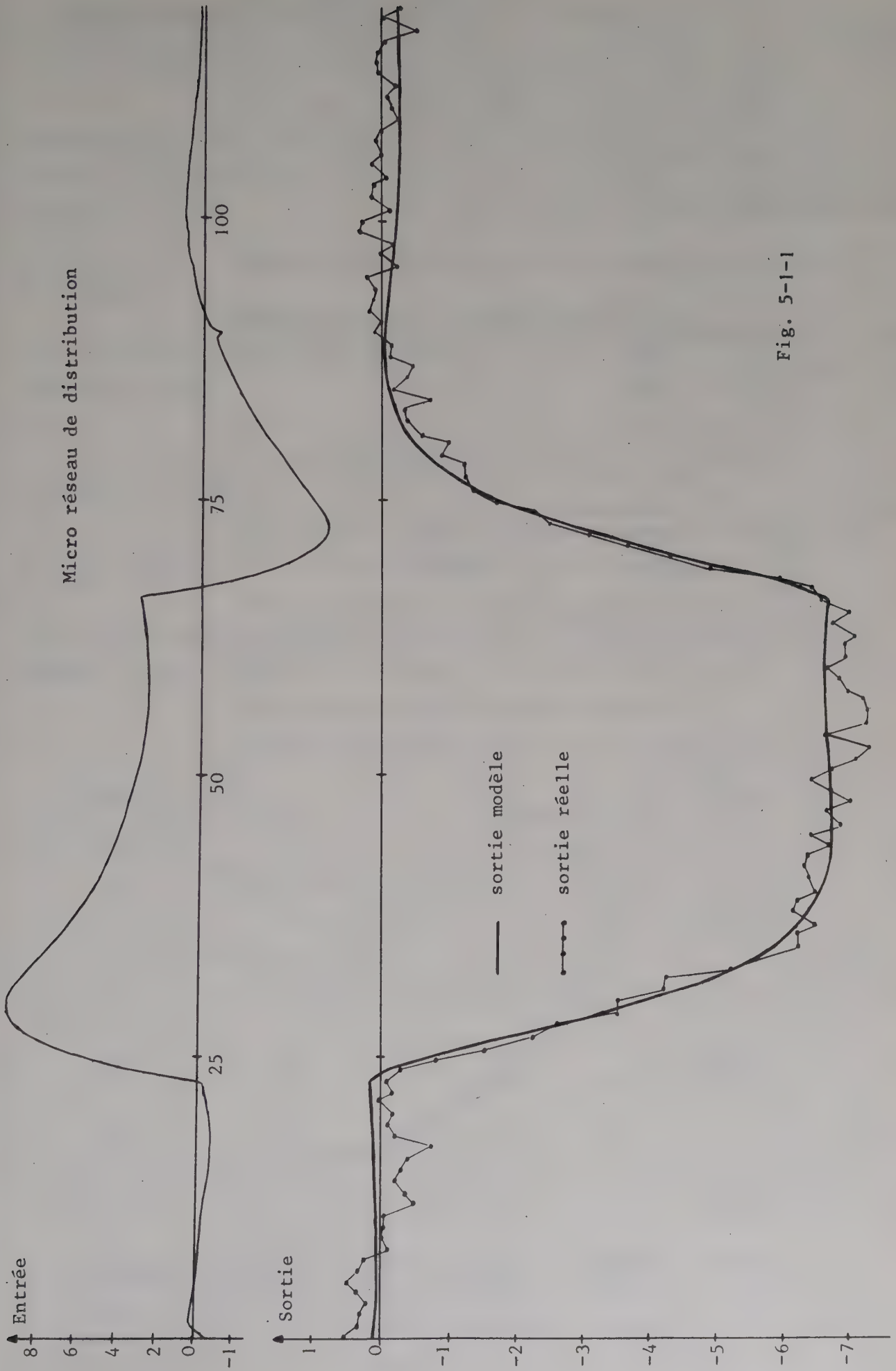


Fig. 5-1-1

Nous adoptons le filtre simplifié pour l'estimation des variances :

$$\hat{\phi}'_{t+1} = \hat{\phi}'_t + \frac{k_t}{2} (\hat{\delta}'_t^2 + \hat{M}'_{\delta\delta t} - \hat{\phi}'_t)$$

$$k_{t+1} = k_t - \frac{k_t^2}{2}$$

qu'on initialise avec $\hat{\phi}'_0 = (\hat{\sigma}'_0)^2 = (0,3)^2$

$$k_0 = 1$$

L'estimation finale par F.N.L.A. est alors :

$\hat{a}_T = -0,96964$	$\hat{b}_T = -0,08259$	$\hat{\sigma}'_T = 0,3466$
$\hat{\sigma}_{aT} = 0,00051$	$\hat{\sigma}_{bT} = 0,00130$	
(coefficient de corrélation : $\rho_{ab} = -0,610$)		

Nous comparons avec les résultats du maximum de vraisemblance

(M.V.) :

$\hat{a} = -0,96939$	$\hat{b} = -0,08296$	$\hat{\sigma}' = 0,3415$
$\sigma_a = 0,00055$	$\sigma_b = 0,00095$	
(coefficient de corrélation : $\rho_{ab} = -0,701$)		

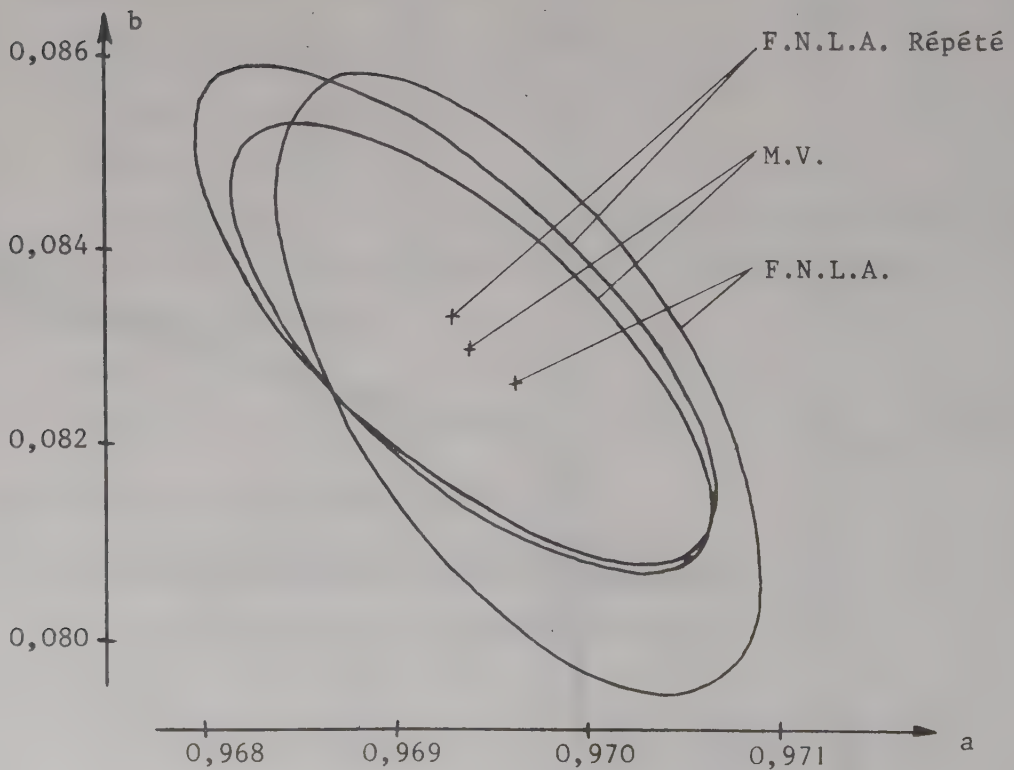


Fig. 5-1-2

Nous pouvons apprécier la concordance de ces résultats sur la figure 5-1-2 où l'on a représenté dans le plan paramétrique (a,b) les estimations accompagnées de l'ellipsoïde de confiance à 95 % correspondant.

Notons que les variances finales étant très inférieures aux variances initiales, l'élimination finale des connaissances à priori a très peu d'effet et les variations n'apparaissent pas sur les chiffres significatifs retenus.

b) Influence de l'initialisation des variances des paramètres et du bruit

Bien que les connaissances à priori précédemment introduites soient raisonnables, nous majorons maintenant très largement l'incertitude initiale en prenant :

$$\begin{aligned}\hat{a}_0 &= 0 & \hat{b}_0 &= 0 \\ \hat{\sigma}_{a0} &= 10 & \hat{\sigma}_{b0} &= 10\end{aligned}$$

D'autre part, nous allons faire deux séries d'expériences en initialisant σ' par les valeurs extrêmes $\hat{\sigma}'_0 = 0,1$ et $\hat{\sigma}'_0 = 1$.

Après élimination des connaissances à priori sur les estimations finales, nous reprenons celles-ci comme initialisation pour un deuxième passage des mêmes données, puis de même pour un troisième passage.

Nous obtenons ainsi les résultats suivants :

— 1ère série de passages avec $\hat{\sigma}'_0 = 0,1$ (voir tableau I)

Tableau I	$\hat{a}$	σ_a	$\hat{b}$	σ_b	$\hat{\sigma}'$	ρ_{ab}
Initialisation	0	10	0	10	0,1	0
avant élimination	-0,96994	0,00054	-0,08423	0,00148	0,36133	
après élimination	-0,96994	0,00054	-0,08423	0,00148		-0,580
Initialisation	"	"	"	"	"	0
avant élimination	-0,96954	0,00036	-0,08322	0,00077	0,35116	
après élimination	-0,96931	0,00055	-0,08328	0,00104		-0,709
Initialisation	"	"	"	"	"	0
avant élimination	-0,96931	0,00036	-0,08327	0,00068	0,34926	
après élimination	-0,96932	0,00055	-0,08326	0,00104		-0,706

—2ème série de passages avec $\hat{\sigma}'_0 = 1$ (voir tableau II)

Tableau II	$\hat{a}$	$\hat{\sigma}_a$	$\hat{b}$	$\hat{\sigma}_b$	$\hat{\sigma}'$	$\hat{\rho}_{ab}$
Initialisation	0	10	0	10	1	0
avant élimination	-0,96868	0,00129	-0,07969	0,00446	0,66919	
après élimination	-0,96868	0,00129	-0,07969	0,00446		-0,201
Initialisation	"	"	"	"	"	0
avant élimination	-0,96928	0,0052	-0,08318	0,00102	0,35942	
après élimination	-0,96932	0,0058	-0,08323	0,00110		-0,701
Initialisation	"	"	"	"	"	0
avant élimination	-0,96932	0,00037	-0,08325	0,00070	0,34940	
après élimination	-0,96932	0,00055	-0,08326	0,00104		-0,706

Nous observons que dès le premier passage, les estimations obtenues ont un sens. On peut prévoir qu'avec un enregistrement plus long, même l'influence de l'initialisation du bruit deviendrait finalement négligeable. D'autre part, en suivant la procédure "hors ligne" ci-dessus (initialisation, élimination finale, réinitialisation, etc...), les résultats convergent très rapidement vers des valeurs indépendantes de l'initialisation.

Ces derniers résultats, reportés sur la figure 5-1-2 s'avèrent plus proches du maximum de vraisemblance que le filtrage non répété, mais dans la mesure où celui-ci est déjà satisfaisant, il apparaît que le point de vue "hors ligne" ne doit être adopté que si les connaissances à priori sont réellement très pauvres.

A partir des trois estimations récapitulées sur la figure 5-1-2, nous avons calculé les "sortie-modèle" correspondantes. Celles-ci sont indiscernables à l'échelle de la figure 5-1-1 où nous donnons seulement la réponse tracée à partir de l'estimation du maximum de vraisemblance.

5-2 IDENTIFICATION D'UNE BATTERIE D'EXTRACTION LIQUIDE - LIQUIDE (Pechiney - Saint Gobain - La Rochelle) -

5-2-1 Présentation du problème

Le système étudié est une batterie d'extraction de terres rares par échange de cations entre une solution aqueuse d'un sel de terres rares et un solvant organique non miscible, du type échangeur cationique. L'installation fonctionne en continu, la grandeur d'action (entrée u) est un débit de reflux. La sortie (s) est la mesure fournie par un spectro-photomètre d'absorption, sensible à la concentration d'un produit clef dans un étage de la batterie.

L'intérêt de cet exemple réside dans le niveau de bruit apparaissant sur l'enregistrement disponible, qui est représenté sur la figure 5-2-1.

On constate d'abord qu'en début d'enregistrement, la sortie est loin d'être constante alors que l'entrée est nulle.

D'autre part, en fin d'enregistrement, alors que l'entrée est revenue à zéro, la sortie ne se stabilise pas à un niveau voisin des premières valeurs, alors que physiquement la transmittance entrée-sortie est stable.

Il est donc clair que la sortie observée est constituée non seulement du signal résultant de la commande, mais d'une perturbation de grande amplitude, qui est visiblement loin d'être une séquence blanche et pour laquelle il faut définir un modèle convenable.

Nous admettons que la transmittance du processus est du 2ème ordre, en retenant le modèle suivant :

$$\begin{cases} x_t = -a_1 x_{t-1} - a_2 x_{t-2} + b_0 u_t + b_1 u_{t-1} \\ d_t = d_{t-1} + \delta_t \\ s_t = x_t + d_t \end{cases}$$

L'équation du bruit d_t définit celui-ci comme une *dérivée*, représentée par le modèle processus de Wiener.

Nous prenons à priori la valeur $\Phi = 1$ comme variance de δ , valeur qui paraît vraisemblable et que nous ne cherchons pas à identifier par la suite.

Au total, nous estimons :

- l'état d_t de la perturbation
- les quatre paramètres a_1 a_2 b_0 b_1
- les deux états correspondant à la transmittance :

$$H(Z) = \frac{b_0 + b_1 Z^{-1}}{1 + a_1 Z^{-1} + a_2 Z^{-2}}$$

La dimension de l'état généralisé est donc de 7.

5-2-2 Résultats de l'identification

$$\begin{array}{ll} \text{— Initialisation choisie : } \hat{a}_1 = 0 & \sigma = 2 \\ & \hat{a}_2 = 0 \quad \sigma = 2 \\ & \hat{b}_0 = 0 \quad \sigma = 1 \\ & \hat{b}_1 = 0 \quad \sigma = 1 \end{array}$$

— Une identification répétée avec élimination des connaissances à priori donne finalement :

$$\begin{array}{ll} \hat{a}_1 = -1,4611 & \sigma_{a1} = 0,0515 \\ \hat{a}_2 = 0,6133 & \sigma_{a2} = 0,0471 \\ \hat{b}_0 = 0,0129 & \sigma_{b0} = 0,0067 \\ \hat{b}_1 = 0,0364 & \sigma_{b1} = 0,0093 \end{array}$$

(notons que l'intervalle de confiance à priori recouvrait largement ces valeurs)

— Comparaison [sortie objet - sortie modèle] :

Nous traçons la réponse de $\hat{H}(z)$ à la séquence d'entrée (fig. 5-2-2). En raison de la présence de la dérive, il est impossible de superposer correctement la sortie calculée à la sortie observée, même en choisissant une composante continue convenable.

Nous nous trouvons dans un cas où, en raison du bruit fortement corréllé, la minimisation d'une norme [sortie objet - sortie modèle] ne peut pas donner de résultat convenable.

5-2-3 Comparaison avec l'identification directe de la séquence de pondération

Le même enregistrement a été utilisé pour calculer directement la séquence de pondération du processus (voir chapitre 6).

Nous représentons sur la figure 5-2-3 l'original de $\hat{H}(Z)$ et la séquence de pondération identifiée par la méthode décrite en 6-2(cf. Collin 1970).

L'écart entre ces deux résultats apparaît compatible avec la pauvreté en information de l'enregistrement.

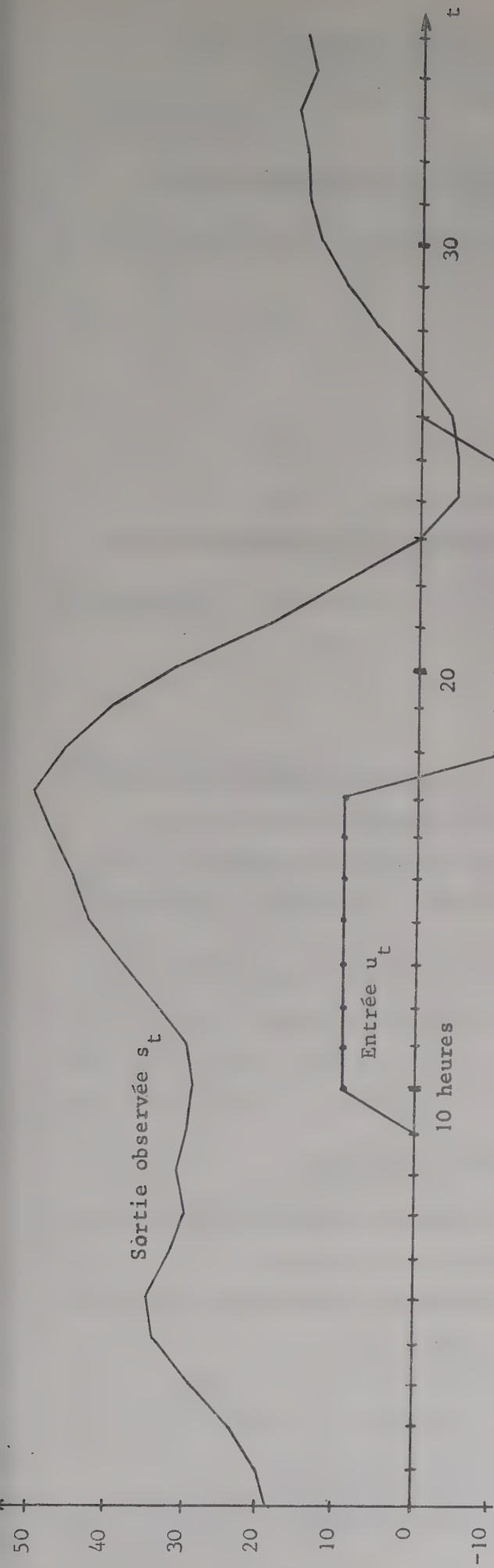


Fig. 5-2-1

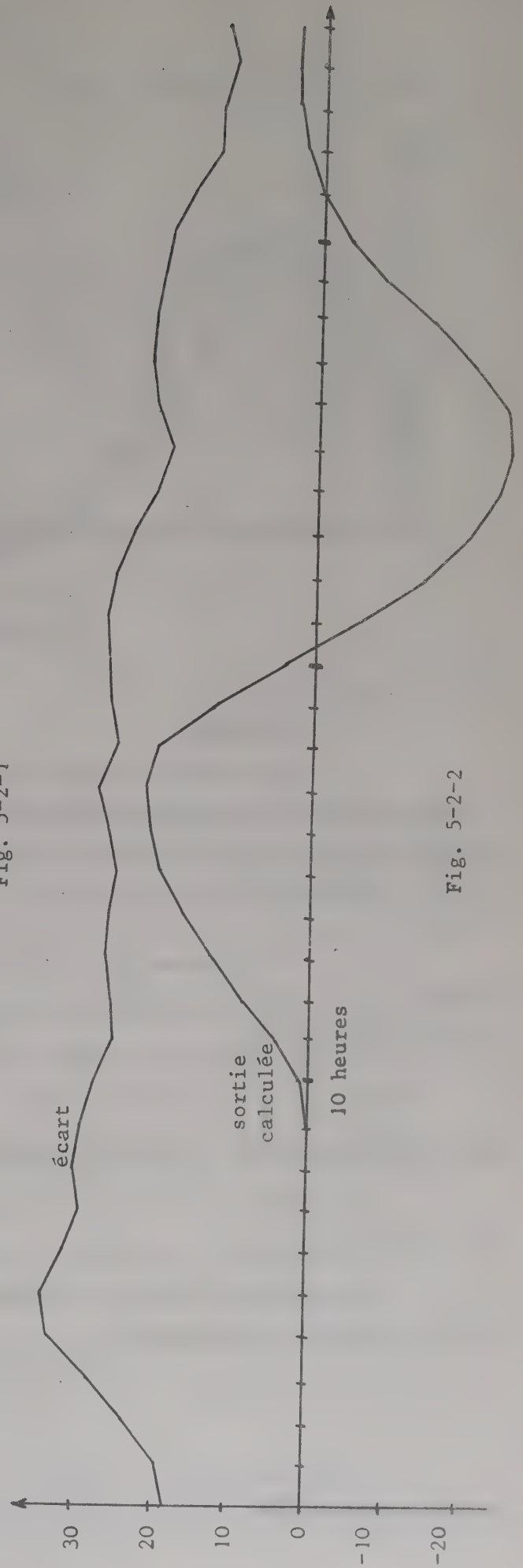


Fig. 5-2-2

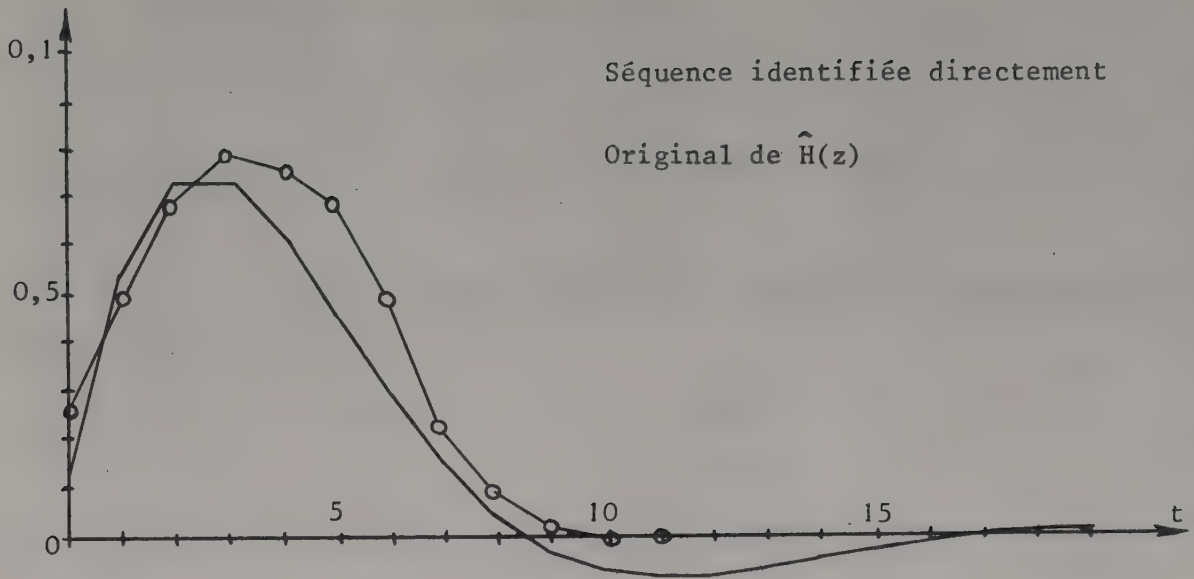


Fig. 5-2-3

5-2-4 Remarques

La première tentative d'identification effectuée à partir de l'enregistrement ci-dessus avait été conduite en effectuant un traitement préalable des données: On relève les écarts de la sortie par rapport à une droite ajustée graphiquement, ce qui revient à faire une première estimation du bruit.

Cette étape, qui n'est d'ailleurs pas toujours possible, est évitée avec la méthode utilisée ici et qui peut se généraliser à bien d'autres types de perturbation.

5-3 IDENTIFICATION DE LA CHAÎNE DE TANGAGE D'UN HELICOPTERE -

L'intérêt de cet exemple est qu'il s'agit d'un système instable avec état initial inconnu.

Les enregistrements sur lesquels nous avons travaillé nous ont été communiqués par le DERA-GERBIOS*.

* Anciennement CERA.

5-3-1 Présentation du problème

La transmittance [angle d'assiette/commande au manche] est caractérisée par :

$$H(s) = \frac{A + Bs}{1 + Cs + Ds^2 + Es^3}$$

d'où la représentation d'état :

$$\dot{\underline{y}}(t) = \begin{bmatrix} 0 & 0 & -a_0 \\ 1 & 0 & -a_1 \\ 0 & 1 & -a_2 \end{bmatrix} \underline{y}(t) + \begin{bmatrix} b_0 \\ b_1 \\ 0 \end{bmatrix} u(t)$$

$$s(t) = [0 \ 0 \ 1] \underline{y}(t)$$

$$\begin{aligned} a_0 &= 1/E \\ a_1 &= C/E \\ a_2 &= D/E \\ b_0 &= A/E \\ b_1 &= B/E \end{aligned}$$

$u(t)$: commande au manche

$s(t)$: angle d'assiette.

Discrétisation : Approximation du premier ordre :

$$y_{t+1} = y_t + T(A y_t + B u_t)$$

$$s_t = C y_t \quad (T, \text{période d'échantillonnage} = 0,1 \text{ s})$$

Données de l'identification :

Nous disposons de trois enregistrements (Essais I, II et III). Cette circonstance est défavorable car, l'état initial étant inconnu, il faudra réinitialiser l'état pour chaque nouvel essai. Par contre, seule la première initialisation des paramètres est à déterminer : le résultat final d'un essai est naturellement pris comme initialisation de l'essai suivant.

De plus, ces essais ont été effectués dans des circonstances différentes. Nous avons apprécié les écarts types des erreurs de mesure au vu des enregistrements : Essais I et II : 1 pour l'entrée, 2 pour la sortie
Essai III : 0,1 pour l'entrée, 0,1 pour la sortie.

5-3-2 Résultats expérimentaux

Nous avons d'abord pris une estimation initiale relativement large, soit 0 pour les 5 paramètres, accompagnés d'écarts types égaux à 20.

Nous avons estimé graphiquement les valeurs de $u(0)$, $s(0)$ et leurs dérivées sur les trois enregistrements, en accompagnant ces estimations d'une variance convenable.

Théoriquement, pour un estimateur strictement optimal, le résultat final ne dépendrait pas de l'ordre dans lequel sont utilisés les trois

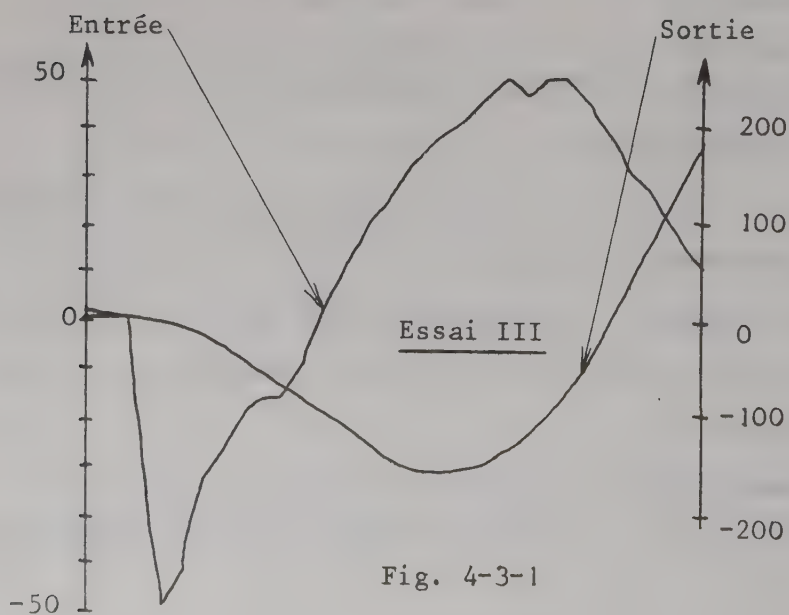
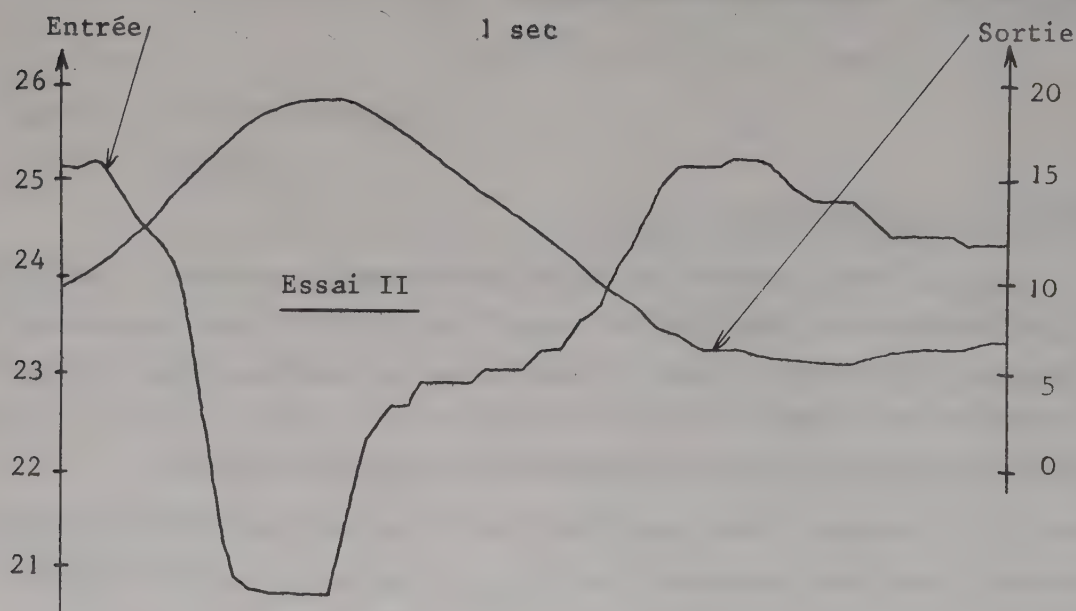
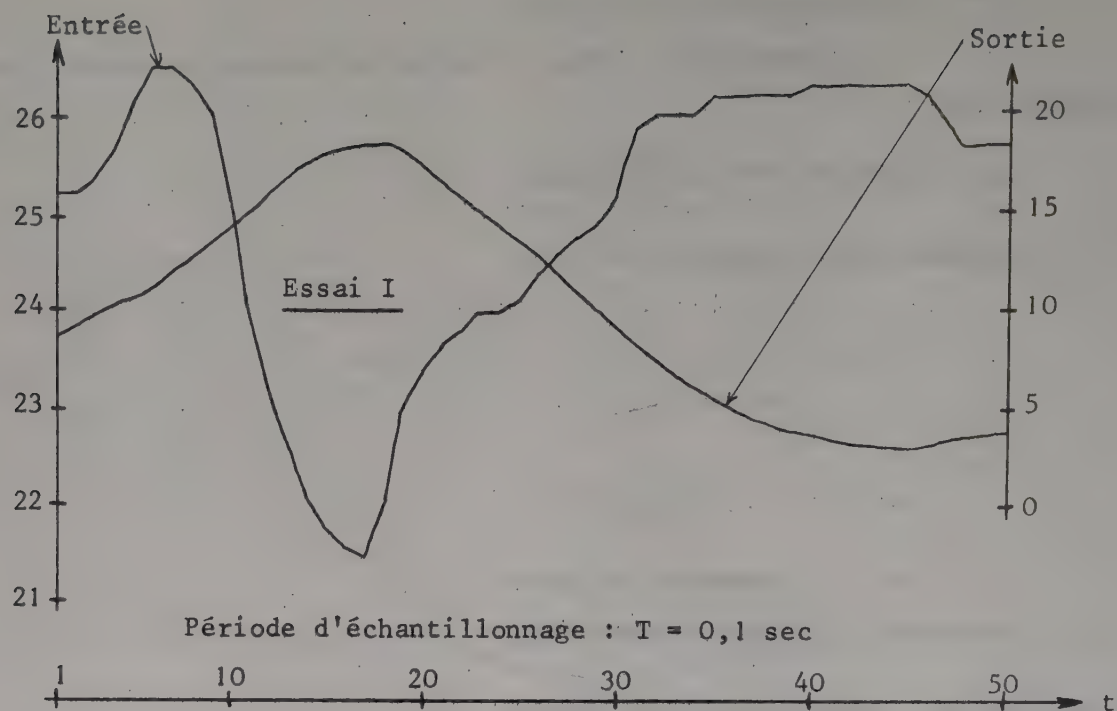


Fig. 4-3-1

enregistrements. C'est pourquoi on a pris les enregistrements dans des ordres différents afin de comparer les résultats, d'où le tableau ci-dessous :

	b_0	b_1	a_1	a_2	a_3
Initialisation	0	0	0	0	0
Essai II	0,24	5,11	0,93	-1,15	1,43
Essai I	0,01	6,27	0,23	-0,13	0,70
Essai III	0,47	10,38	1,06	-1,26	1,10
Ecart type final	0,036	0,152	0,076	0,062	0,048
Initialisation	0	0	0	0	0
Essai I	0,17	7,45	0,53	-0,28	0,83
Essai III	0,35	10,01	0,95	-0,88	1,00
Essai II	0,28	10,33	0,69	-0,98	0,93
Ecart type final	0,041	0,180	0,077	0,078	0,054
Initialisation	0	0	0	0	0
Essai III	35,85	4,41	6,90	1,16	5,93
Essai II	-0,08	8,47	-0,068	-0,58	0,47
Essai I	-0,01	7,91	-0,05	-0,48	0,51
Ecart type final	0,038	0,025	0,078	0,092	0,053

Les résultats des deux premières séries d'essais sont à peu près cohérents. Par contre, dans la troisième série, on trouve des résultats aberrants qui montrent justement que la stabilité de l'algorithme n'est pas garantie.

Nous utilisons alors les résultats des deux premières séries pour initialiser à nouveau l'algorithme :

$$\hat{b}_0 = 0,40, \quad \hat{\sigma} = 0,2$$

$$\hat{b}_1 = 10,00, \quad \hat{\sigma} = 1,0$$

$$\hat{a}_1 = 0,75, \quad \hat{\sigma} = 0,4$$

$$\hat{a}_2 = -1,10, \quad \hat{\sigma} = 0,4$$

$$\hat{a}_3 = 1,00, \quad \hat{\sigma} = 0,4$$

Trois nouvelles séries d'essais donnent les résultats suivants :

	b_0	b_1	a_1	a_2	a_3
Initialisation	0,40	10,00	0,75	-1,10	1,00
Essai III	0,42	8,68	0,98	-0,38	0,84
Essai I	0,55	10,01	1,23	-1,24	1,15
Essai II	0,29	9,81	0,66	-1,02	0,85
Ecart type final	0,035	0,159	0,070	0,061	0,046
Initialisation	0,40	10,00	0,75	-1,10	1,00
Essai II	0,06	7,58	0,31	-0,15	0,81
Essai III	0,29	10,39	0,72	-0,95	0,96
Essai I	0,30	10,08	0,70	-1,02	0,89
Ecart type final	0,032	0,159	0,065	0,064	0,043
Initialisation	0,40	10,00	0,75	-1,10	1,00
Essai I	0,23	7,90	0,60	-0,53	0,84
Essai III	0,54	10,13	1,23	-1,20	1,16
Essai II	0,31	10,02	0,72	-1,04	0,90
Ecart type final	0,034	0,158	0,067	0,090	0,045

Nous constatons que, dans chaque série, les résultats intermédiaires diffèrent un peu, mais tous les résultats finals sont très voisins les uns des autres.

D'autre part, les écarts types obtenus sont suffisamment inférieurs aux écarts types primitifs (rapport supérieur à 5) pour que l'élimination finale des connaissances à priori soit inutile.

En conclusion, nous comparerons nos résultats avec ceux obtenus par le DERA par la méthode du modèle :

DERA	ENSM
A = 0,495	$\hat{b}_0/\hat{a}_0 = 0,426$
B = 13,5	$\hat{b}_1/\hat{a}_0 = 14,45$
C = -0,75	$\hat{a}_1/\hat{a}_0 = -1,46$
D = 0,9	$\hat{a}_2/\hat{a}_0 = 1,29$
E = 1,5	$1/\hat{a}_0 = 1,43$

5-4 IDENTIFICATION D'UN TURBO REACTEUR (ATAR 9 C DE LA SNECMA) -

L'intérêt de ce problème est qu'il s'agit d'un système non linéaire.

Nous identifions la chaîne [débit de carburant - régime] d'un turbo réacteur.

Le couple d'entrée-sortie $[u(t), s(t)]$ est représenté sur la figure 5-4-1.

On note :

- $u(t) = C(t)$ = débit carburant (litres/heure)
- $s(t) = N(t)$ = régime (tours/minute)

5-4-1 Caractérisation du processus

Nous adoptons des équations de la forme générale suivante (équation d'état du 1er ordre non linéaire)

$$\dot{x} = f(x, u)$$

$$s = x \text{ (+ erreurs de mesure)}$$

La fonction $f(x, u)$ est non linéaire et doit elle-même être caractérisée. Nous supposons à priori une expression polynomiale :

$$f(u, x) = \alpha_1 + \alpha_2 x + \alpha_3 u + \alpha_4 x^2 + \alpha_5 u x + \alpha_6 u^2 + \alpha_7 x^3 + \dots$$

Certains de ces coefficients peuvent cependant être considérés comme nuls à partir des résultats de l'analyse physico-mathématique du système ; les équations théoriques permettent en effet de tracer les courbes $f(x, u)$ à u constant et à x constant, et on constate que les premières sont sensiblement linéaires, les secondes ne l'étant absolument pas. Ceci suggère d'éliminer tous les termes en $u^2, u^3, u^4, \dots$, dans le développement ci-dessus, soit finalement :

$$\dot{x} = f(u, x) = \theta_1 + \theta_2 x + \theta_3 u + \theta_4 u x + \theta_5 x^2 + \theta_6 u x^2 + \theta_7 x^3 + \dots$$

5-4-2 Résultats de l'identification

Nous avons considéré différents modèles de complexité croissante, depuis le modèle linéaire à 3 paramètres :

$$\dot{x} = \theta_1 + \theta_2 x + \theta_3 u$$

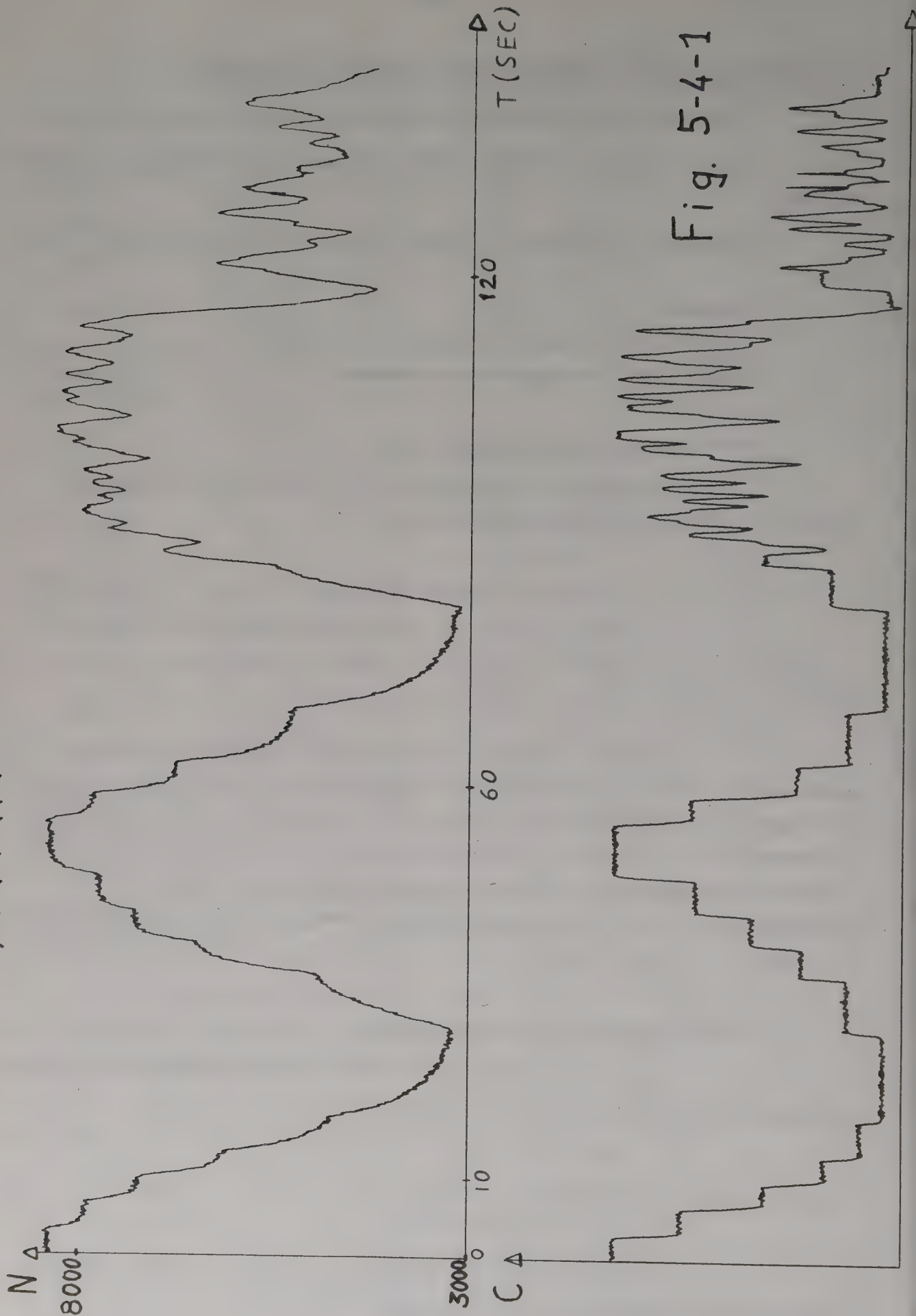
jusqu'au modèle à 7 paramètres :

$$\dot{x} = \theta_1 + \dots + \theta_7 x^3$$

L'identification est effectuée hors ligne, en deux passages successifs avec élimination des connaissances à priori.

L'équation d'état continue est discrétisée par approximation

ATAR 9C



du premier ordre.

Pour chaque modèle, nous comparons la sortie objet et la sortie modèle et nous calculons un indicateur d'écart :

$$I = \sqrt{\frac{\sum (\text{sortie objet} - \text{sortie modèle})^2}{\sum (\text{sortie objet} - \text{val. moyenne})^2}}$$

(I sera exprimé en %)

- Modèle linéaire :

$$\dot{x} = \theta_1 + \theta_2 x + \theta_3 u$$

Indicateur I = 31,01 %

Comparaison sortie-objet, sortie-modèle : fig. 5-4-2

- Modèle à 4 paramètres :

$$\dot{x} = \theta_1 + \theta_2 x + \theta_3 u + \theta_4 u x$$

I = 8,83 %

fig. 5-4-3

- Modèle à 6 paramètres :

$$\dot{x} = \theta_1 + \theta_2 x + \theta_3 u + \theta_4 u x + \theta_5 x^2 + \theta_6 u x^2$$

I = 3,94 %

fig. 5-4-4

- Modèle à 7 paramètres :

$$\dot{x} = \theta_1 + \theta_2 x + \theta_3 u + \theta_4 u x + \theta_5 x^2 + \theta_6 u x^2 + \theta_7 x^3$$

I = 3,95 %

fig. 5-4-5

Au delà de 6 paramètres, l'indicateur d'écart ne décroît plus. Au contraire, on constate une légère augmentation (de 3,94 % à 3,95 %) qui nous paraît sans signification, l'identification ayant été faite à partir d'unités réduites et des erreurs d'arrondi ont pu être commises lors de la conversion des diverses grandeurs et du calcul final de l'indicateur.

- Modèle retenu :

On a recherché finalement, parmi les modèles à 5 paramètres, la combinaison pour laquelle l'indicateur d'écart est le plus faible.

Selon qu'en plus des termes linéaires, on retient les termes en $(x u \text{ et } u^2)$, $(x^2 \text{ et } u x)$, $(x^2 \text{ et } u^2)$, $(x^2 \text{ et } x^3)$, on trouve respectivement pour I : 8,4 %, 7,8 %, 7,1 %, 5,1 %.

SIMULATION ATAR 9 C

ISO-SECTION TUYERE 3600 PLAGE REGIME 2974 / 8425

CRITERE QUADRATIQUE = 31.01

ECAAT MAX. = 11003 TR/MN

Modèle :

$$\dot{x} = \theta_1 + \theta_2 x + \theta_3 u$$

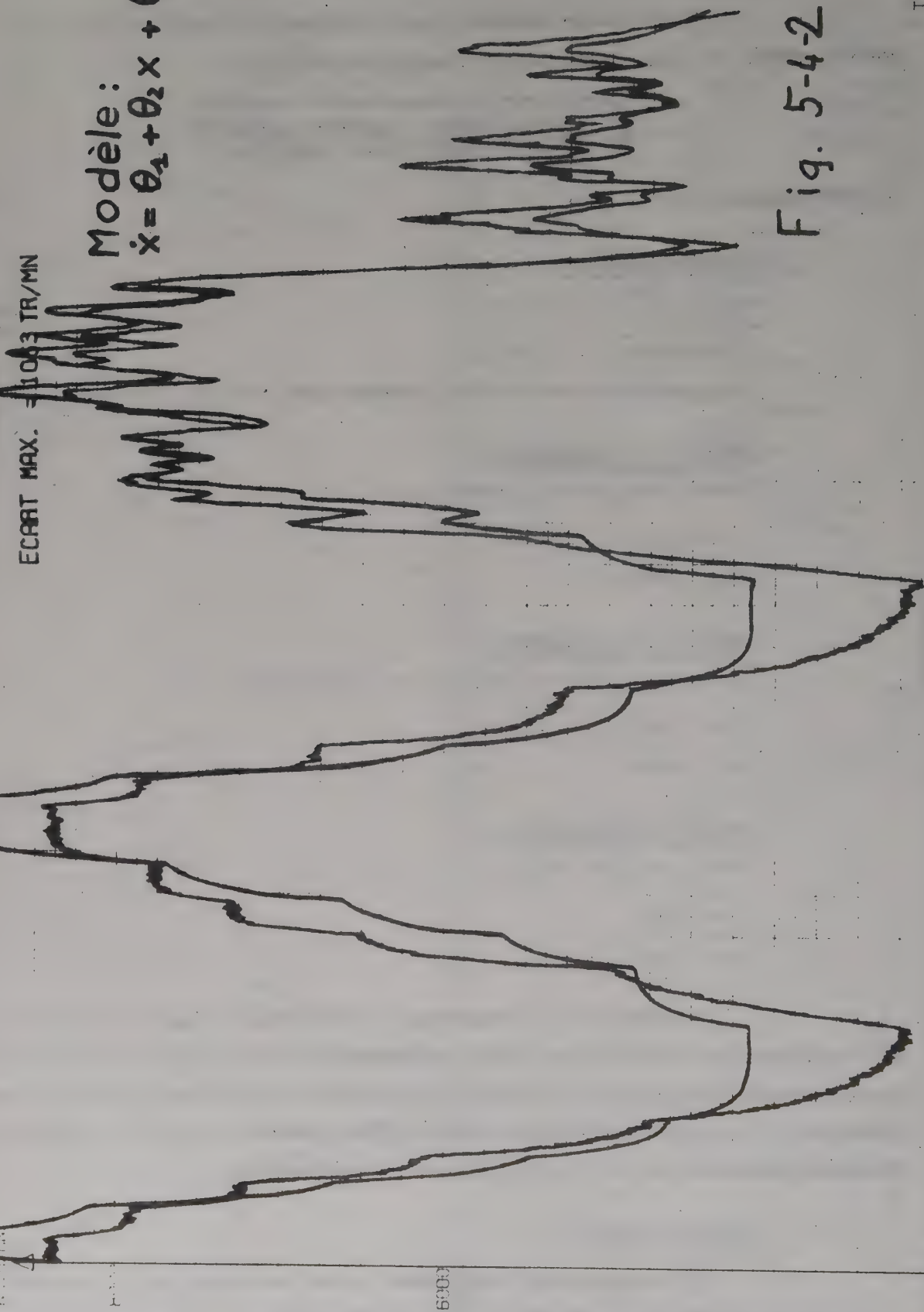


Fig. 5-4-2

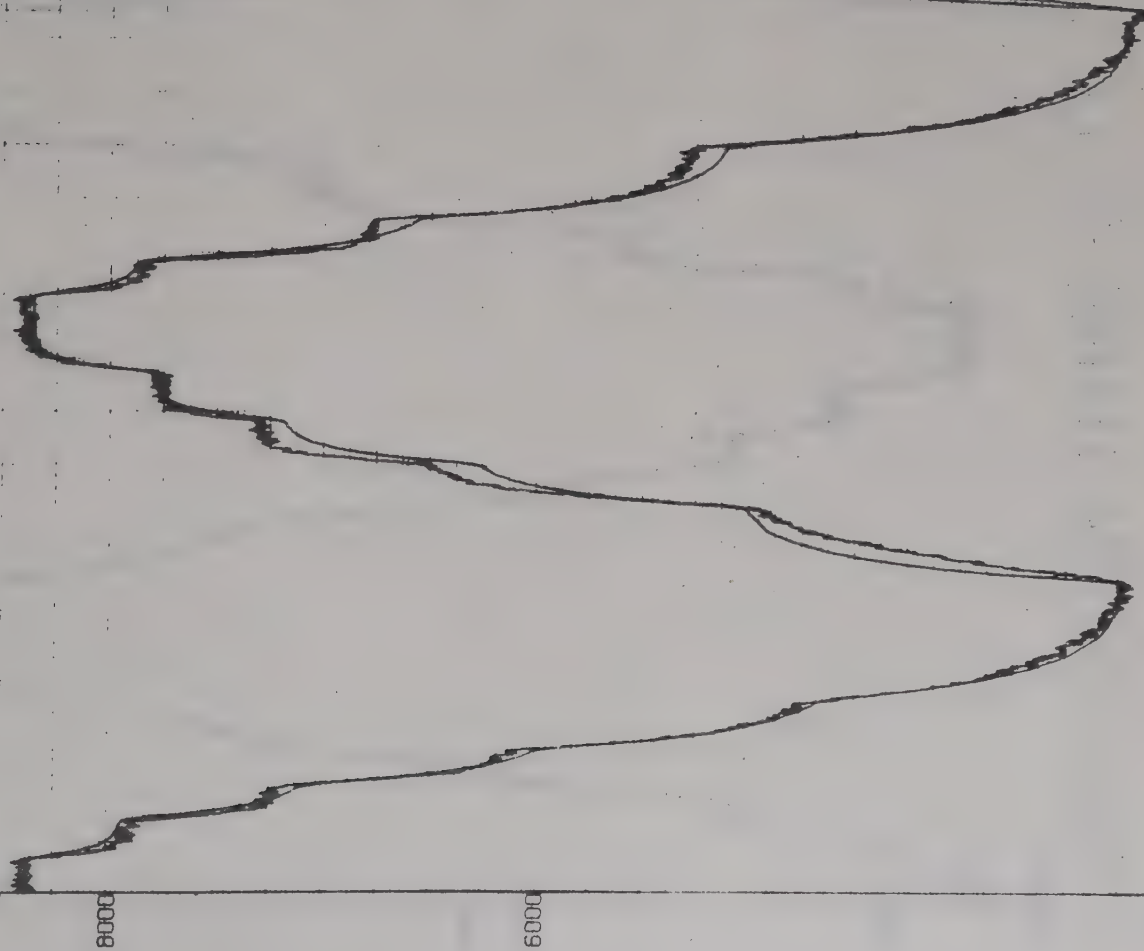
SIMULATION ATAR 9 C

ISO-SECTION TUYERE 3600 PLAGE REGIME 2974 / 8425

CRITERE QUADRATIQUE = 8.83

ECART MAX. = 638 TR/MIN

N (TR/MIN)



Modèle:

$$\dot{x} = \theta_1 + \theta_2 x + \theta_3 u + \theta_4 u x$$

Fig. 5-4-3

T (SEC)

50

SIMULATION ATAR 9 C

N (TR/MN)



ISO-SECTION TUYERE 3600 PLAGE REGIME 2974 / 8425

CRITERE QUADRATIQUE = 3.94

ECART MAX. = 241 TR/MN

Modèle :

$$\dot{x} = \theta_1 + \theta_2 x + \theta_3 u + \theta_4 ux + \theta_5 x^2 + \theta_6 ux^2$$

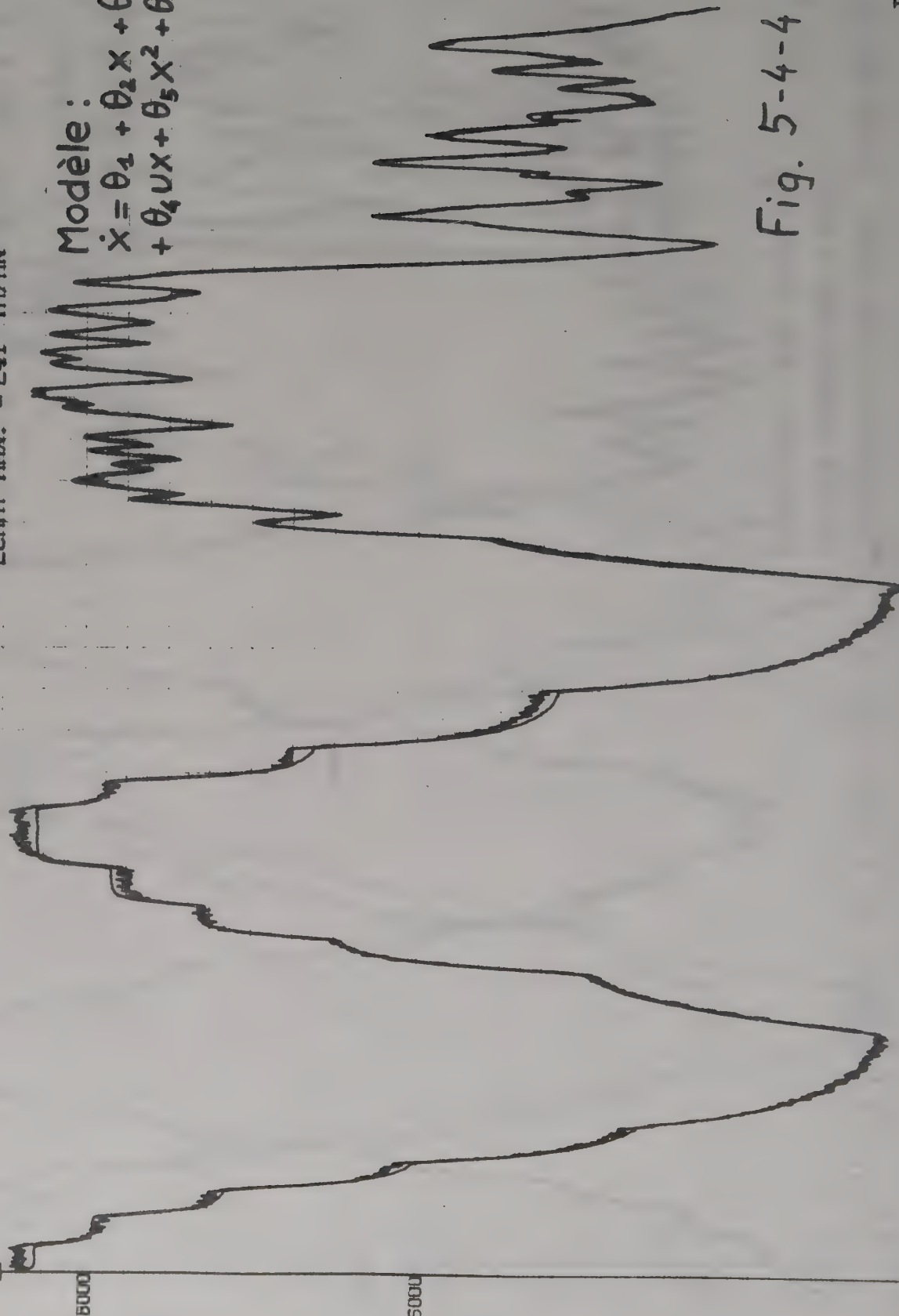
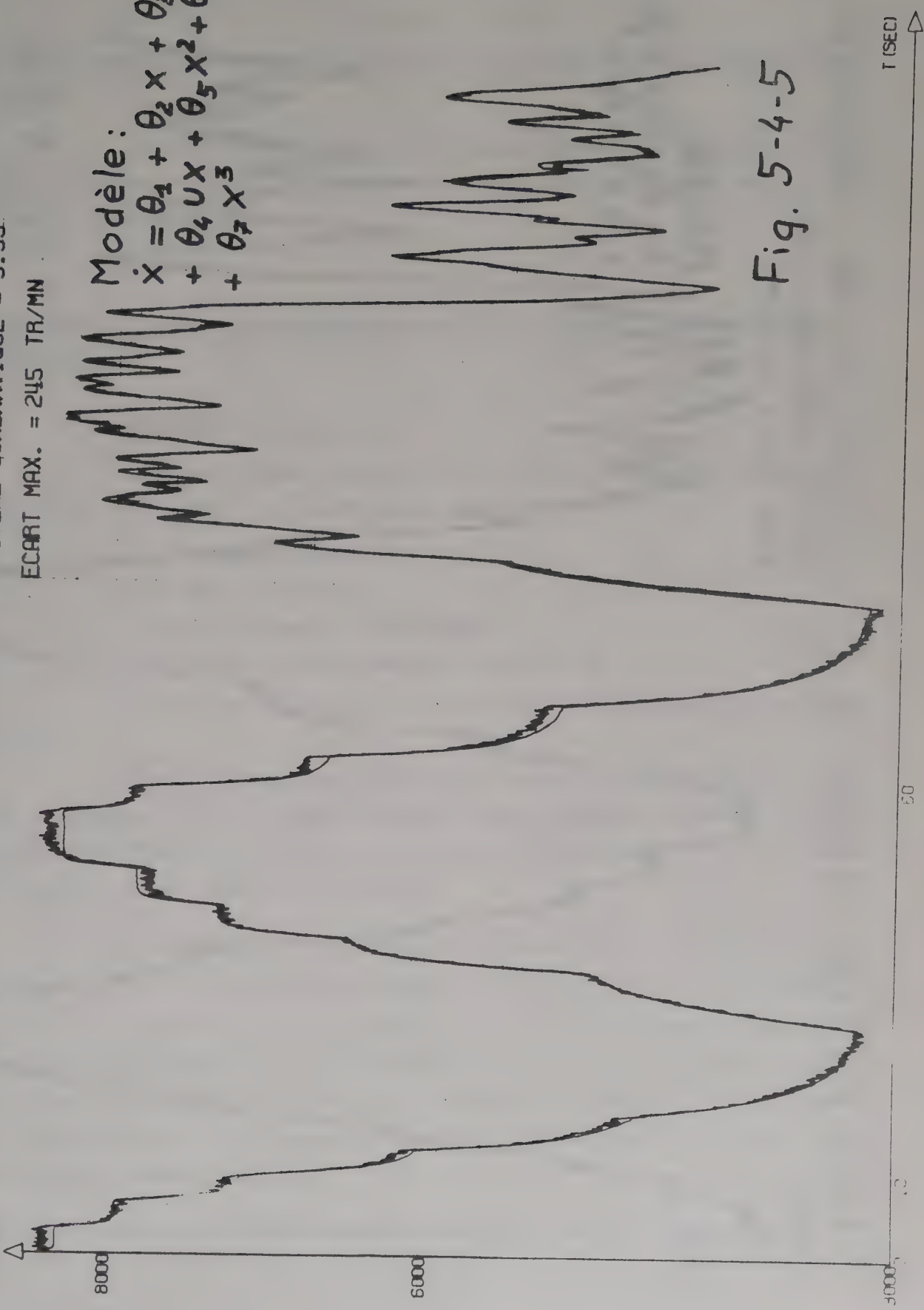


Fig. 5-4-4

SIMULATION ATAR 90

ISO-SECTION TUYERE 3600 PLAGE REGIME 2974 / 8425
 CRITERE QUADRATIQUE = 3.95
 ECART MAX. = 245 TR/MN



Modèle:

$$\dot{X} = \theta_1 + \theta_2 X + \theta_3 U$$

$$+ \theta_4 U X + \theta_5 X^2 + \theta_6 U X^2$$

$$+ \theta_7 X^3$$

Fig. 5-4-5

SIMULATION ATAR 9 C

N (TR/MN)

Δ

ISO-SECTION TUYERE 3600 PLAGE REGIME 2974 / 8425

CRITERE QUADRATIQUE = 5.33

ECART MAX. = 379 TR/MN

Modèle:

$$\dot{x} = \theta_4 + \theta_2 x + \theta_3 u + \theta_5 x^2 + \theta_7 x^3$$

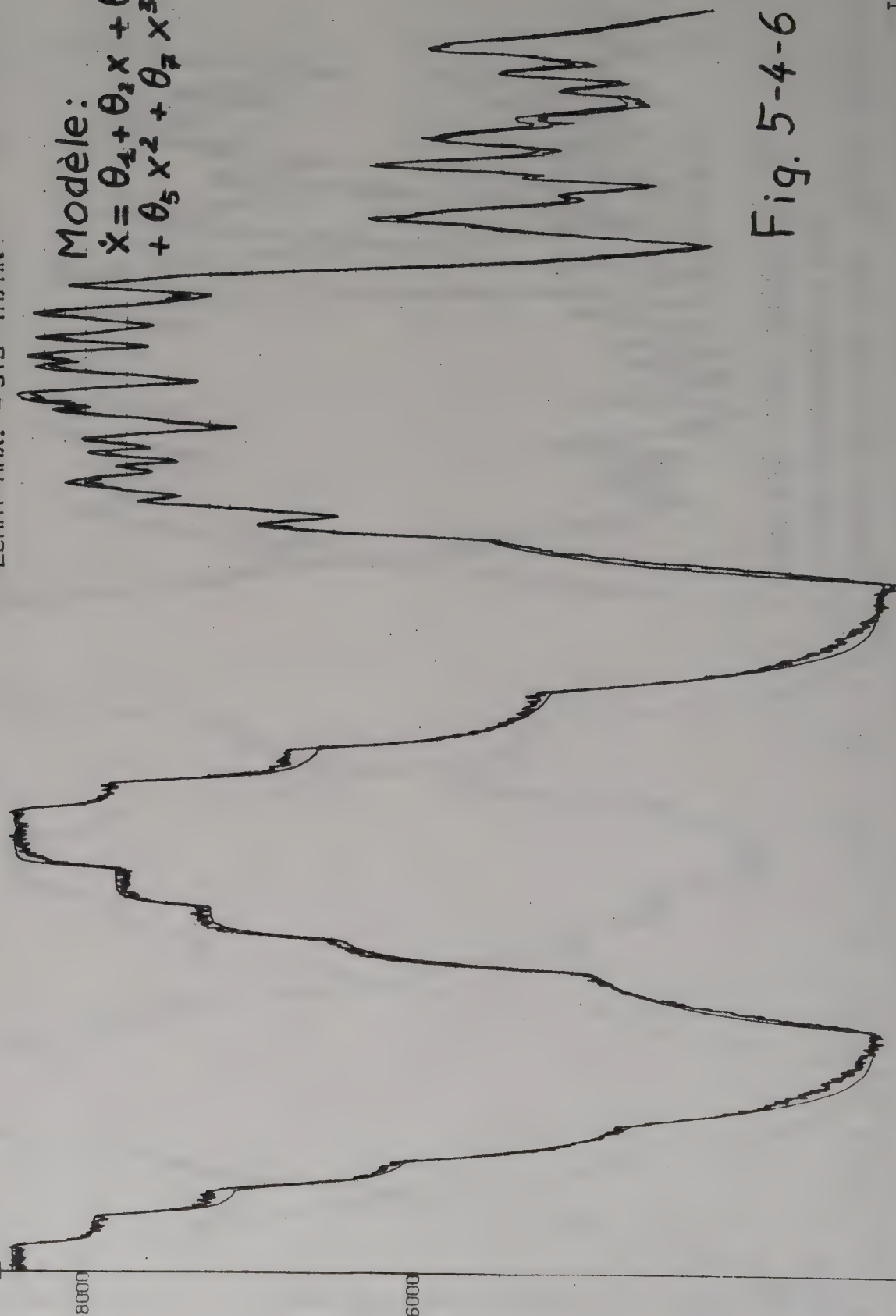


Fig. 5-4-6

On retiendra donc le modèle :

$$\dot{x} = \theta_1 + \theta_2 x + \theta_3 u + \theta_5 x_2 + \theta_7 x_3$$

qui offre un compromis convenable complexité - précision (sorties objet-modèle : fig. 5-4-6).

Ce modèle se justifie d'ailleurs en constatant que les courbes $f(u, x)$ à x constant sont non seulement approximativement linéaires, mais encore pratiquement parallèles.

5-5 IDENTIFICATION D'UNE UNITE PILOTE GENERATEUR DE VAPEUR (IDENTIFICATION MULTIVARIABLE) -

5-5-1 Description du processus

Nous nous sommes intéressés à la chaîne de température de l'unité pilote du laboratoire d'Automatique de l'E.N.S.M.

Les grandeurs de commande sont :

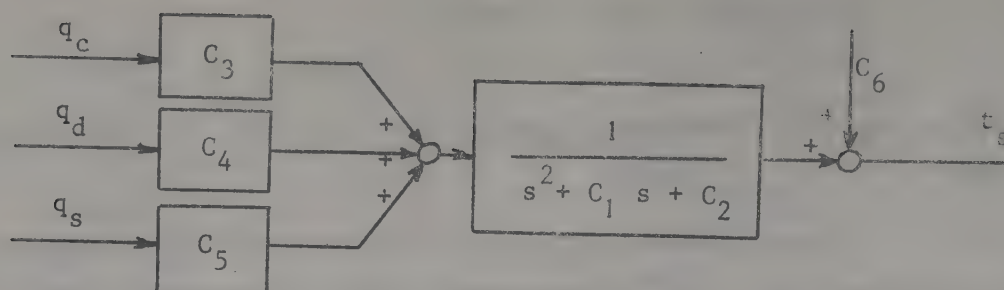
- le débit de fuel q_c
- le débit d'eau de désurchauffe q_d
- le débit de vapeur admission turbine q_s

La grandeur de sortie considérée est la température de vapeur surchauffée t_s .

D'autres commandes, telles que le débit d'eau d'alimentation (asservi au niveau ballon), la dépression au foyer (asservie au débit de fuel), sont l'objet de régulations locales, et ce que nous identifions est en fait le générateur muni de ses régulations annexes.

D'autre part, la température est mesurée par l'intermédiaire d'un capteur que nous considérons également comme faisant partie du processus.

Il se trouve d'ailleurs que le capteur qui était utilisé lors de nos expériences présentait une constante de temps assez importante, de l'ordre de celles du processus lui-même, ce qui masquait les différences de comportement dynamique pour les différentes entrées. C'est pourquoi nous avons pu adopter alors le modèle suivant :



- 1'unité de temps est 1'heure
- q_c , q_d , q_s exprimés en kg/heure de fuel, d'eau de désurchauffe et de vapeur
- t_s en degrés C.

Les coefficients C_1 à C_6 sont identifiés à partir de l'échantillon d'entrée-sortie de la figure 5-5-1.

Le pas d'échantillonnage est de 10 secondes. Celui-ci est suffisamment petit pour que la discrétisation des équations différentielles par l'approximation du 1er ordre soit suffisante.

5-5-2 Résultats de l'identification

Nous présentons l'évolution des estimations $\hat{C}_1$ à $\hat{C}_6$ obtenues lors de deux passages de l'enregistrement de la figure 5-5-1, à partir d'initialisations différentes, l'une en partant d'une bonne connaissance a priori du processus, l'autre à partir de valeurs initiales toutes nulles. La concordance des résultats peut être appréciée sur la figure 5-5-2.

5-5-3 Identification en temps réel

Nous avons pu mettre en oeuvre une identification en temps réel de la transmittance Débit de désurchauffe - Température, les autres commandes (fuel et vapeur) étant maintenues constantes.

Les résultats détaillés seront présentés dans un prochain DEA (Piasco, juin 1971). Disons seulement que, les mesures étant effectuées toutes les dix secondes, le temps de calcul à chaque itération du filtre ne dépasse pas 1 seconde sur le calculateur utilisé (EAI 640).

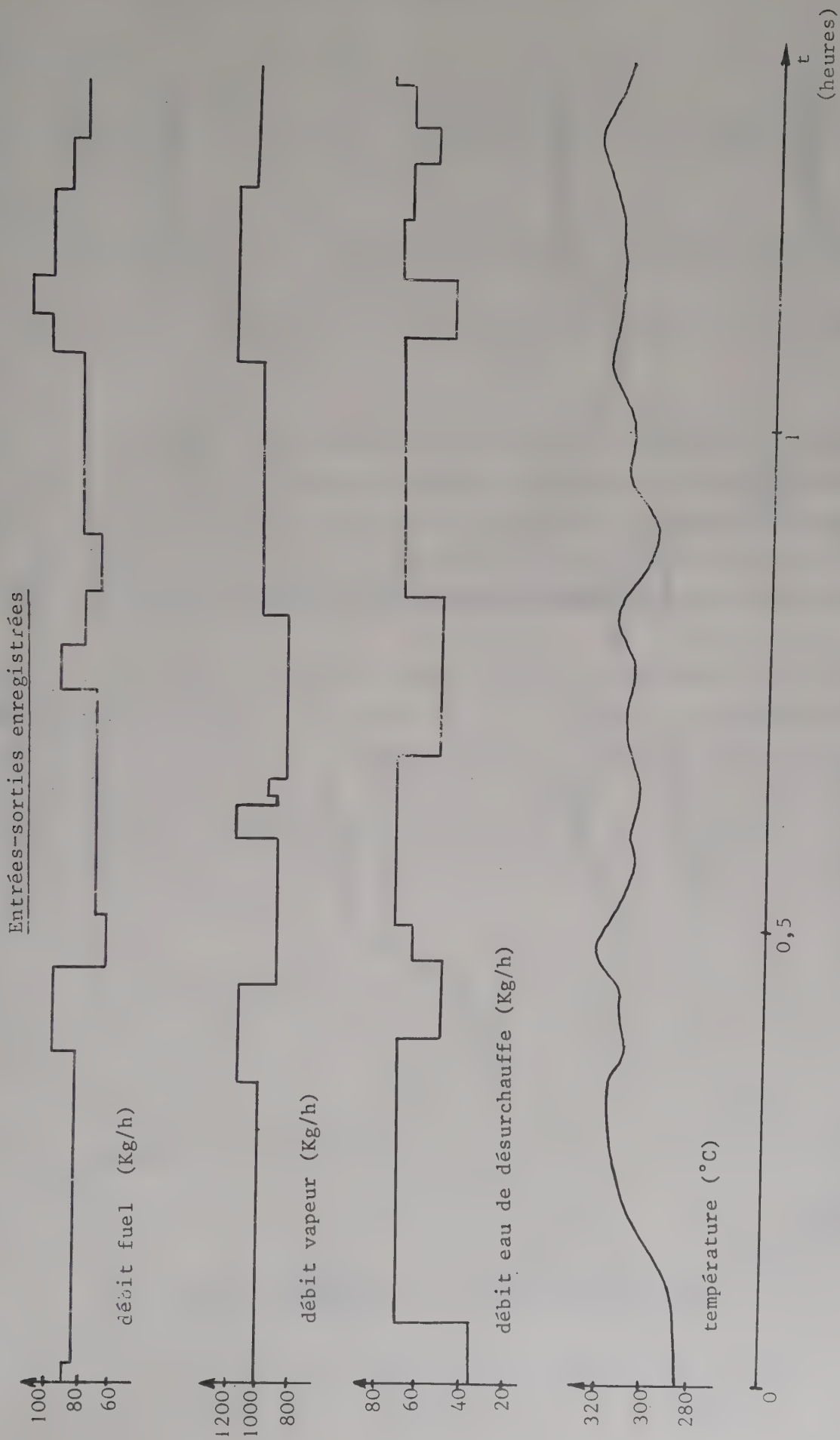
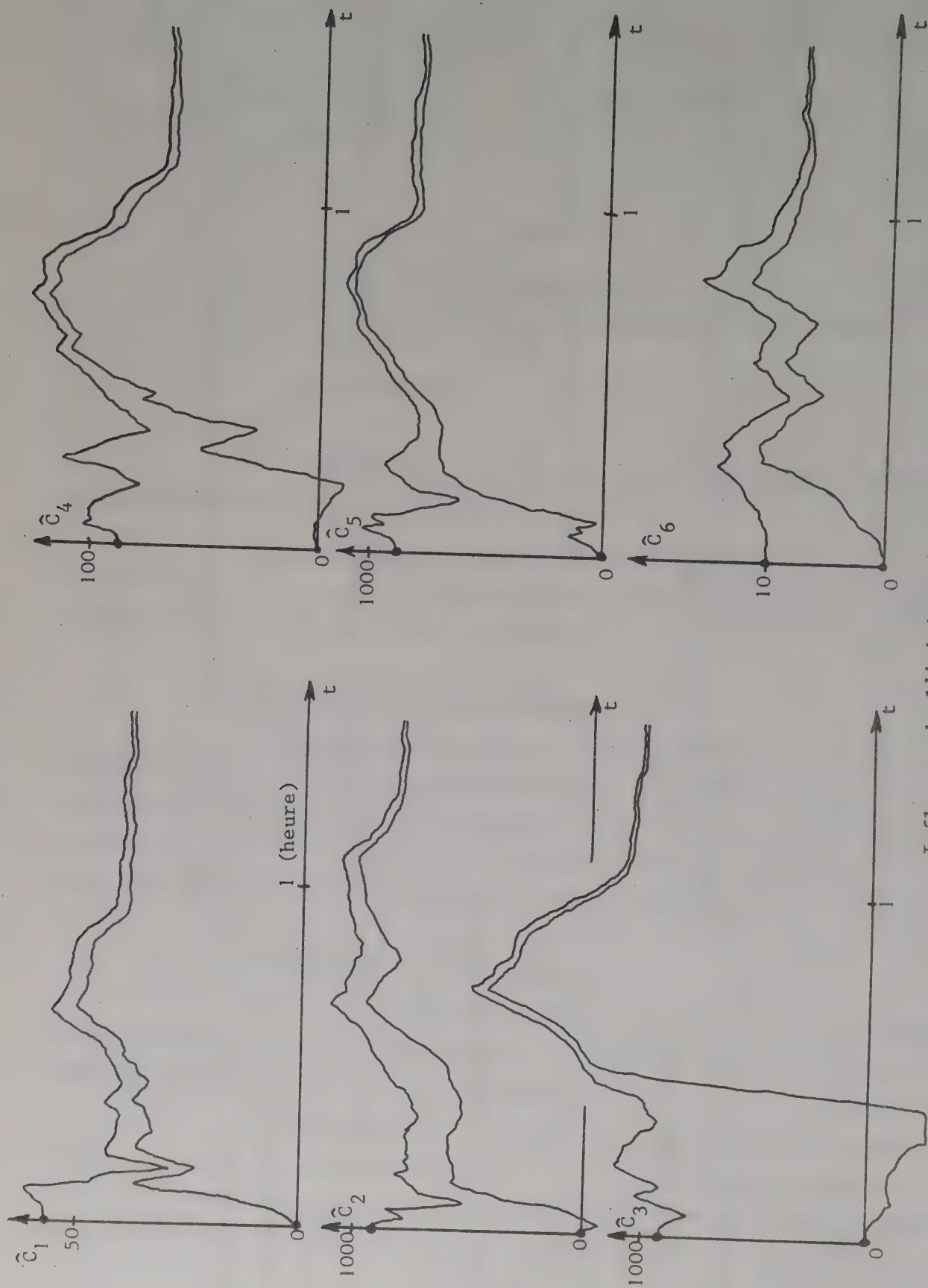


Fig. 5-5-1



Influence de l'initialisation

Fig. 5-5-2

Chapitre 6

APPLICATIONS PARTICULIERES DU FILTRAGE LINEAIRE A L'IDENTIFICATION

6-1 Identification des systèmes linéaires représentés par leur séquence de pondération.

6-2 Identification par fonctions modulatrices et filtrage linéaire.

Bien qu'en règle générale, les problèmes d'identification exigent un traitement non linéaire des informations, nous présentons ici des extensions pour deux méthodes linéaires, de portée moins générale, mais qui ont leur intérêt dans leur domaine d'application : l'identification de séquence de pondération et l'identification par fonctions modulatrices.

L'attrait de la représentation "séquence de pondération", ou réponse impulsionnelle discrète est qu'alors les non linéarités des équations d'état généralisées disparaissent. D'autres difficultés surgissent alors : instabilité numérique, sensibilité au bruit. C'est pour surmonter celles-ci que nous proposons d'utiliser les connaissances à priori usuelles sur les réponses impulsionnelles pour initialiser l'algorithme d'identification et en améliorer les performances.

Un autre procédé n'utilisant que des techniques linéaires est l'identification par fonctions modulatrices (Loeb, 1965). Cette technique consiste à effectuer des transformations intégrales sur les échantillons d'entrée-sortie des processus à identifier. Nous montrons ici comment ces transformations peuvent être mécanisées en temps réel par simple filtrage des entrées-sorties, et nous proposons un traitement des données conduisant à un algorithme d'identification en temps réel par voie purement analogique.

6-1 IDENTIFICATION DES SYSTEMES LINEAIRES REPRESENTES PAR LEUR SEQUENCE DE PONDERATION -

6-1-1 Introduction de la représentation "séquence de pondération"

Une des difficultés fondamentales de l'identification provient du fait qu'en général, les équations d'état généralisées sont non linéaires.

Cette difficulté disparaît cependant dans certains cas.

Considérons la forme générale de l'équation récurrente définissant un processus monovarié :

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \dots + a_n y_{t-n} + b_0 u_t + \dots + b_q u_{t-q}$$

Une forme particulière de cette équation est :

$$y_t = h^0 u_t + h^1 u_{t-1} + \dots + h^\tau u_{t-\tau} + \dots + u_{t-m} \quad 6-1-1$$

où le dénominateur de la transmittance en Z a disparu. Nous notons $\underline{h}$ les coefficients du numérateur :

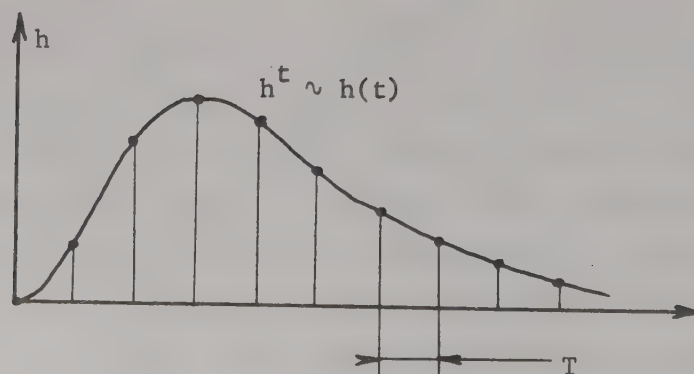
$$\underline{h}^T = [h^0 \ h^1 \ \dots \ h^2 \ \dots \ h^m]$$

h^τ représente la *pondération* de l'entrée u à l'instant $t-\tau$ sur la sortie y à l'instant t .

L'équation (6-1-1) peut être considérée comme l'expression discrète de l'intégrale de convolution pour les systèmes continus :

$$y(t) = \int_0^\infty h(\tau) u(t-\tau) d\tau$$

où $h(t)$ est la réponse impulsionnelle du processus (fig. 6-1)



T = période d'échantillonnage, supposée égale à 1

Fig. 6-1

Nous savons que tout système linéaire stationnaire peut être représenté par sa réponse impulsionnelle (original de la transmittance $H(s)$), ou dans le cas discret par sa séquence de pondération (original de $H(z)$).

Du point de vue de l'identification, l'intérêt majeur de cette représentation provient de ce que, comme nous le verrons, les difficultés liées aux non linéarités de l'équation d'état disparaissent. Un autre intérêt est que, hormis linéarité et stationnarité, aucune hypothèse n'est faite sur la structure des équations du système, et qu'on peut identifier la réponse impulsionnelle d'un processus obéissant par exemple à un système d'équations aux dérivées partielles sans plus de difficulté que pour des équations différentielles ordinaires à constantes localisées.

D'autre part, on peut éventuellement utiliser directement le modèle séquence de pondération pour établir la commande d'un processus (Dawkins 1965, de Larminat 1968, Hock 1968).

En contre partie, les réponses impulsionnelles des processus ont en général une durée m (ou mémoire) théoriquement infinie. Cependant, pour les systèmes asymptotiquement stables, $h(t) \rightarrow 0$ quand $t \rightarrow \infty$, et l'on pourra se limiter pratiquement à m fini. Il n'en reste pas moins que le nombre de paramètres à considérer (dimension de h) sera généralement bien supérieur au nombre de paramètres d'une transformée en Z correspondante. Cette multiplication du nombre de paramètres n'est pas seulement un inconvénient sur le plan de la mécanisation des calculs. D'un certain point de vue, on peut dire que la quantité d'information contenue dans un enregistrement d'entrée-sortie doit alors être répartie sur un plus grand nombre d'inconnues, d'où une détérioration de l'identification. C'est effectivement ce qu'on observe, surtout lorsque la durée de l'échantillon est peu supérieure, ou exactement égale à m .

Il est bien connu que les résultats numériques sont alors extrêmement sensibles aux moindres bruits de mesure, au point de perdre souvent toute signification.

Nous proposons ici une méthode permettant d'améliorer sensiblement les procédures classiques (moindres carrés effectués globalement ou séquentiellement), en introduisant des connaissances à priori dans un algorithme d'identification, sous la forme d'une initialisation convenable.

6-1-2 Mise en équation du problème et initialisation de l'algorithme d'identification

Reprenons l'équation 6-1-1 sous la forme : $y_t = C_t h$
(nous notons $C_t = [u_t \ u_{t-1} \ \dots \ u_{t-m}]$)

Nous pouvons introduire l'équation :

$$\underline{x}_{t+1} = \underline{x}_t (= \underline{h}) \quad 6-1-2$$

D'autre part, en faisant intervenir un bruit de mesure sur la sortie, l'équation (6-5-1) devient :

$$\underline{s}_{t+1} = C_{t+1} \underline{x}_{t+1} + \delta'_{t+1} \quad 6-1-3$$

Nous supposons ici que le bruit de mesure δ'_t est une séquence indépendante de variance Φ' . Si ce bruit est un processus markovien linéaire de caractéristiques connues, il est immédiat de reprendre le problème en augmentant $\underline{x}_t$ de l'état correspondant au modèle du bruit.

Les équations d'état généralisées (6-1-2, 6-1-3) sont ici linéaires et donc l'estimation de $\underline{x}_t = \underline{h}$ ne pose plus de problème théorique.

Les équations d'identification se réduisent alors à :

$$\begin{aligned} \hat{\underline{h}}_t &= \hat{\underline{h}}_{t-1} + \hat{\underline{M}}_{t-1} C_t (C_t \hat{\underline{M}}_{t-1} C_t^T + \Phi')^{-1} (y_t - C_t \hat{\underline{h}}_{t-1}) \\ \hat{\underline{M}}_t &= \hat{\underline{M}}_{t-1} - \hat{\underline{M}}_{t-1} C_t (C_t \hat{\underline{M}}_{t-1} C_t^T + \Phi')^{-1} C_t^T \hat{\underline{M}}_{t-1} \end{aligned}$$

Initialisées par $\hat{\underline{h}}_0 = 0$, $\hat{\underline{M}}_0 = \alpha_0 \mathbf{1}$ (avec $\alpha_0 \rightarrow \infty$), ces équations constituent d'ailleurs une formulation récurrente de la méthode des moindres carrés (Levin 1960) :

$$\hat{\underline{h}}_t \text{ minimise } \sum_{i=1}^t (y_i - C_i \underline{h})^2.$$

Grâce au point de vue statistique, nous pouvons introduire une estimation à priori $\hat{\underline{h}}_0$, ainsi que la matrice $\hat{\underline{M}}_0$ reflétant la confiance accordée à cette estimation.

Plus exactement, soit $\tilde{\underline{h}}_0$ l'écart $\underline{h} - \hat{\underline{h}}_0$. Si nous pouvons estimer à priori que toutes les composantes de cet écart sont inférieures en valeur absolue à une limite supérieure donnée D, nous prendrons une matrice de covariance initiale $\hat{\underline{M}}_0 = S^2 \mathbf{1}$ avec $S = D/2$ (si $\underline{h}_0 \pm D$ est un intervalle de confiance initial à 95,5 %).

Cette initialisation implique cependant que toutes les composantes de $\tilde{\underline{h}}_0$ seraient à priori indépendantes les unes des autres.

Or, les réponses impulsionnelles usuelles (du type exponentiel ou sinus amorti par exemple) sont des fonctions indéfiniment dérivables sur $]0, \infty[$, et l'on choisit évidemment une initialisation $\hat{\underline{h}}_0(t)$ de même type. Par conséquent, $\underline{h}_0(t)$ doit être considéré aussi comme une fonction dérivable.

Pour tenir compte de ce fait dans l'initialisation $\hat{\underline{M}}_0$, nous allons rechercher quelle peut être la fonction d'autocorrélation d'une fonction

aléatoire indéfiniment dérivable.

Considérons que la fonction aléatoire $\tilde{h}_o(t)$ est générée par un bruit blanc $b(t)$ filtré un certain nombre de fois par des transmittances du premier ordre (fig. 6-2).

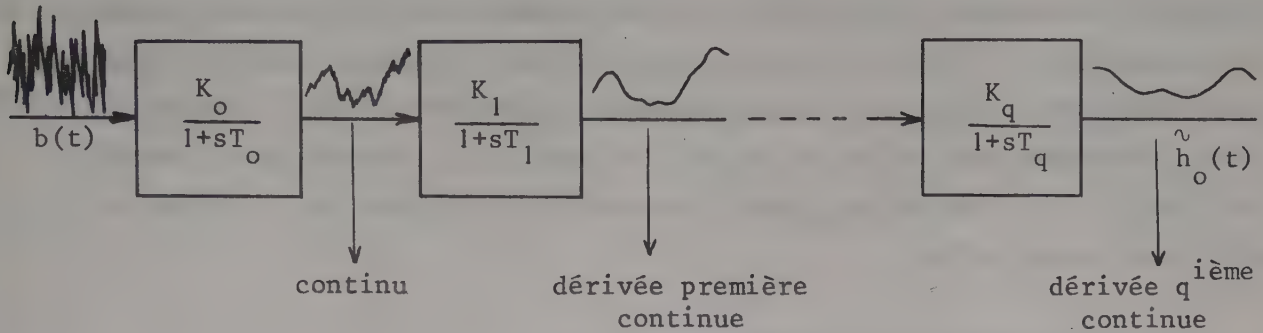


Fig. 6-2

Notons $k_o, k_1, \dots, k_q(t)$ les réponses impulsionnelles des filtres successifs.

On sait que la fonction de covariance $\Phi(\tau) = E(\tilde{h}_o(t) \tilde{h}_o(t+\tau))$ est alors $\Phi(\tau) = k_o(\tau) * k_o(-\tau) * k_1(\tau) * k_1(-\tau) * \dots * k_q(\tau) * k_q(-\tau)$ (* dénote le produit de convolution).

Or, le théorème Centrale Limite nous apprend que, sous des conditions très générales, la fonction résultante d'une série de produits de convolution tend vers une fonction de Gauss. D'où une fonction de covariance de la forme :

$$\Phi(\tau) = S^2 e^{-\left(\frac{\tau}{H}\right)^2}$$

Nous écrirons donc la matrice de covariance :

$$\hat{M}_o = [\hat{M}_o^{ij}] = E[\tilde{h}_o^i \tilde{h}_o^j] = [E(\tilde{h}_o(i) \tilde{h}_o(j))] = [S^2 e^{-\left|\frac{i-j}{H}\right|^2}]$$

Nous pouvons bien entendu raffiner encore l'expression de $\hat{M}_o$ en considérant un écart $\tilde{h}_o$ non stationnaire, pour tenir compte par exemple du fait que $\tilde{h}_o(t) \rightarrow 0$ quand $t \rightarrow \infty$.

Il semble préférable cependant de se limiter à une initialisation de la forme la plus simple possible, et pour cette même raison, on prendra d'ailleurs souvent $\underline{\tilde{h}_o} = \underline{0}$.

L'utilisateur doit alors seulement choisir deux paramètres S et H pour initialiser l'algorithme.

- S est proportionnel, comme nous l'avons vu à l'intervalle de confiance à priori de $\underline{h}$

- H est un "facteur de continuité" attribué à la réponse impulsionnelle. Remarquons que celui-ci est normalement en rapport avec la mémoire m du système, que l'on doit de toute façon se fixer. On peut considérer comme raisonnable des valeurs de H comprises entre m et m/10. Choisi trop grand, H risque d'effacer des variations significatives de la réponse impulsionnelle. Trop petit, M_0 se rapproche alors d'une matrice diagonale et le paramètre H n'a plus alors aucun effet.

6-1-3 Résultats expérimentaux

Nous présentons d'abord les résultats obtenus en simulation à partir des échantillons d'entrée-sortie représentés sur la figure 6-3.

La réponse impulsionnelle est considérée comme nulle pour $t > m = 37$.

Le rapport de l'écart type du bruit de mesure sur l'écart type de la sortie est de 1 %.

A) Nombre de mesures égal à la mémoire du système (fig. 6-4)

L'algorithme est initialisé avec $\underline{h}_0 = \underline{0}$, $M_0 = \alpha I$, α étant suffisamment grand (10^6) pour que les résultats puissent être considérés comme ceux des moindres carrés. Nous présentons le résultat de l'identification pour $t = m = 37$. Nous observons que, malgré le faible rapport bruit/signal, les résultats sont dépourvus de toute signification.

B) Moindres carrés (fig. 6-5)

L'identification est poursuivie jusqu'à l'instant $t = 75$, soit le double de la mémoire du système. Les résultats présentent des "dents de scie" caractéristiques de la méthode des moindres carrés.

C) Identification avec connaissance à priori (fig. 6-6)

Nous avons représenté la courbe exacte, l'estimation initiale, et l'estimation finale obtenue pour $t = 37$, c'est-à-dire que l'on peut faire la comparaison avec les résultats (A).

L'initialisation choisie était la suivante :

$$h_0^i = \frac{1}{3} i e^{-\frac{i}{10}}$$

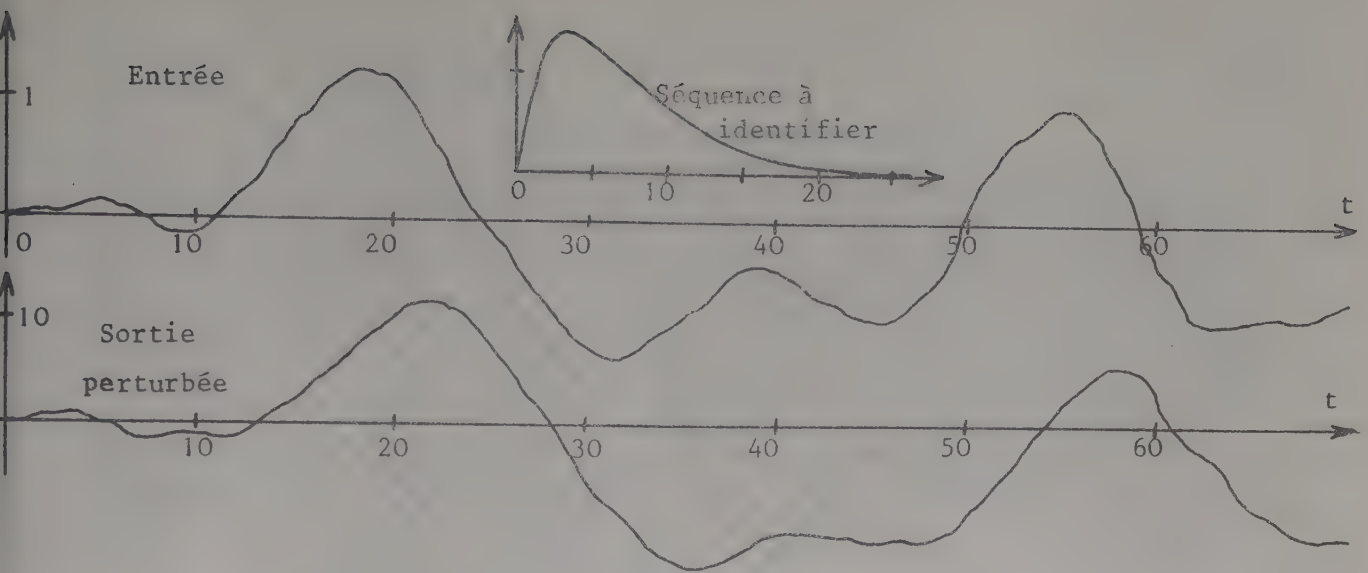


Fig. 6-3

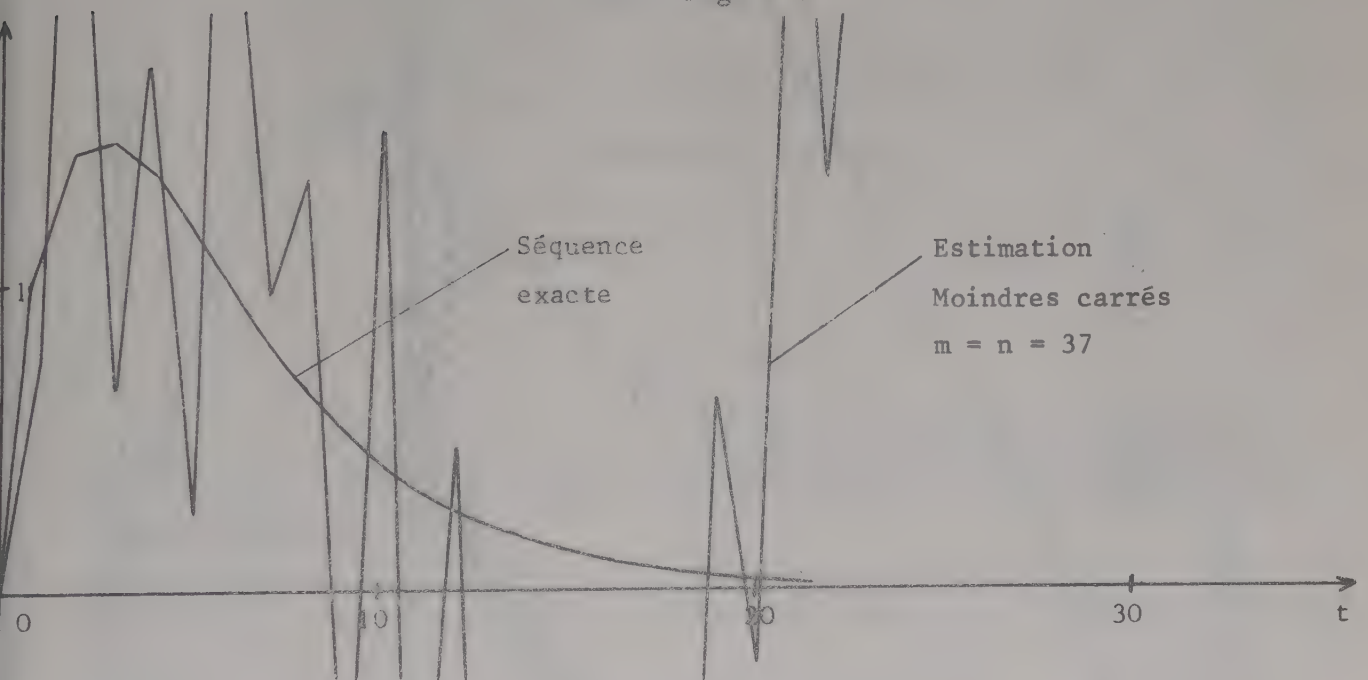


Fig. 6-4

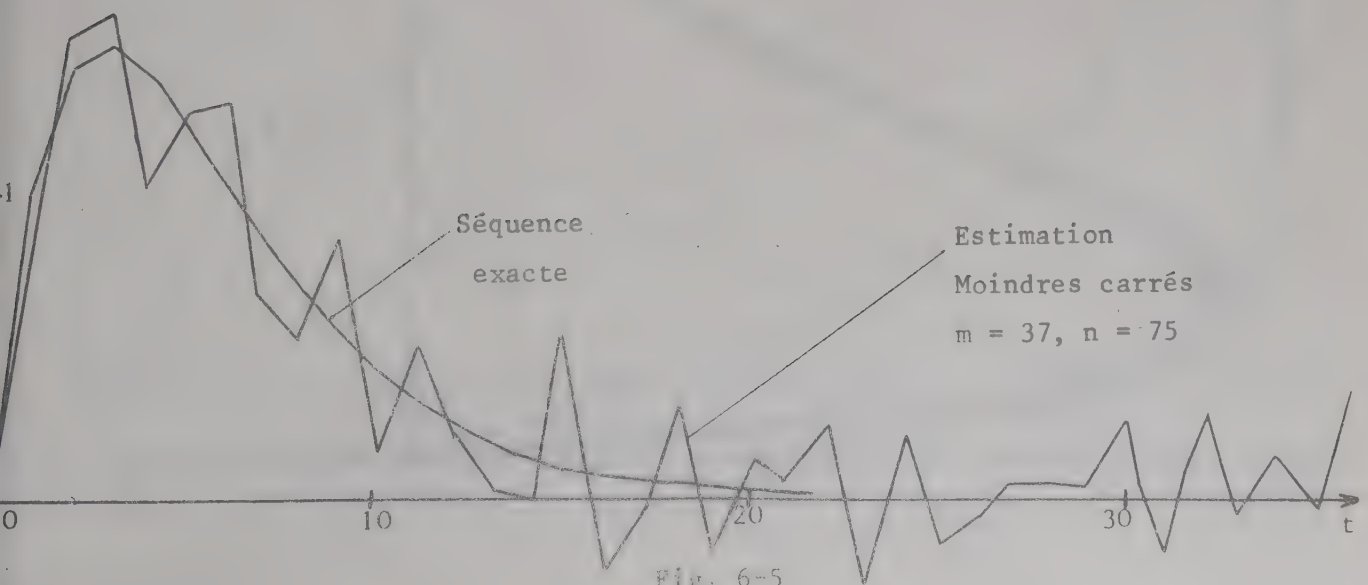


Fig. 6-5

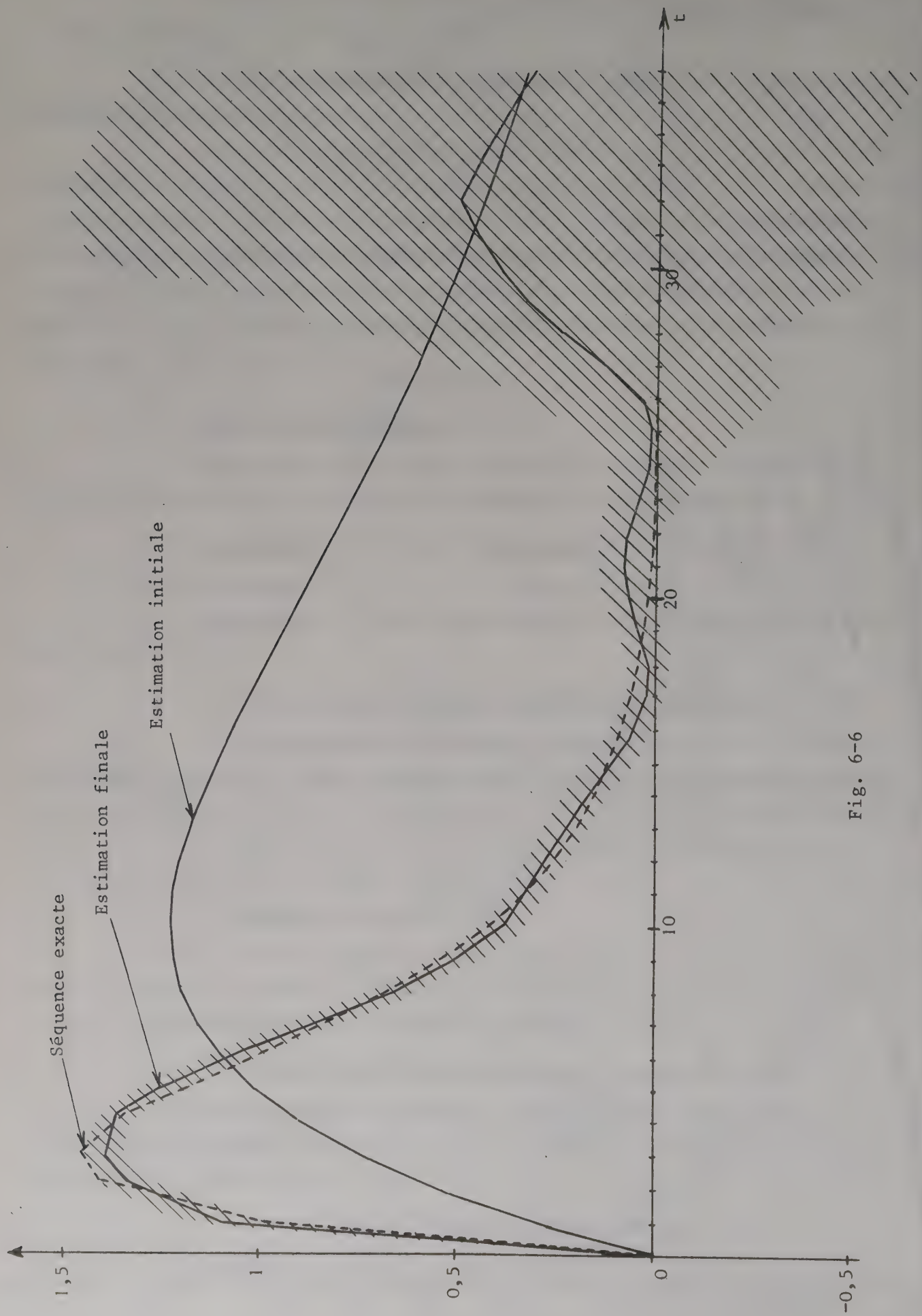


Fig. 6-6

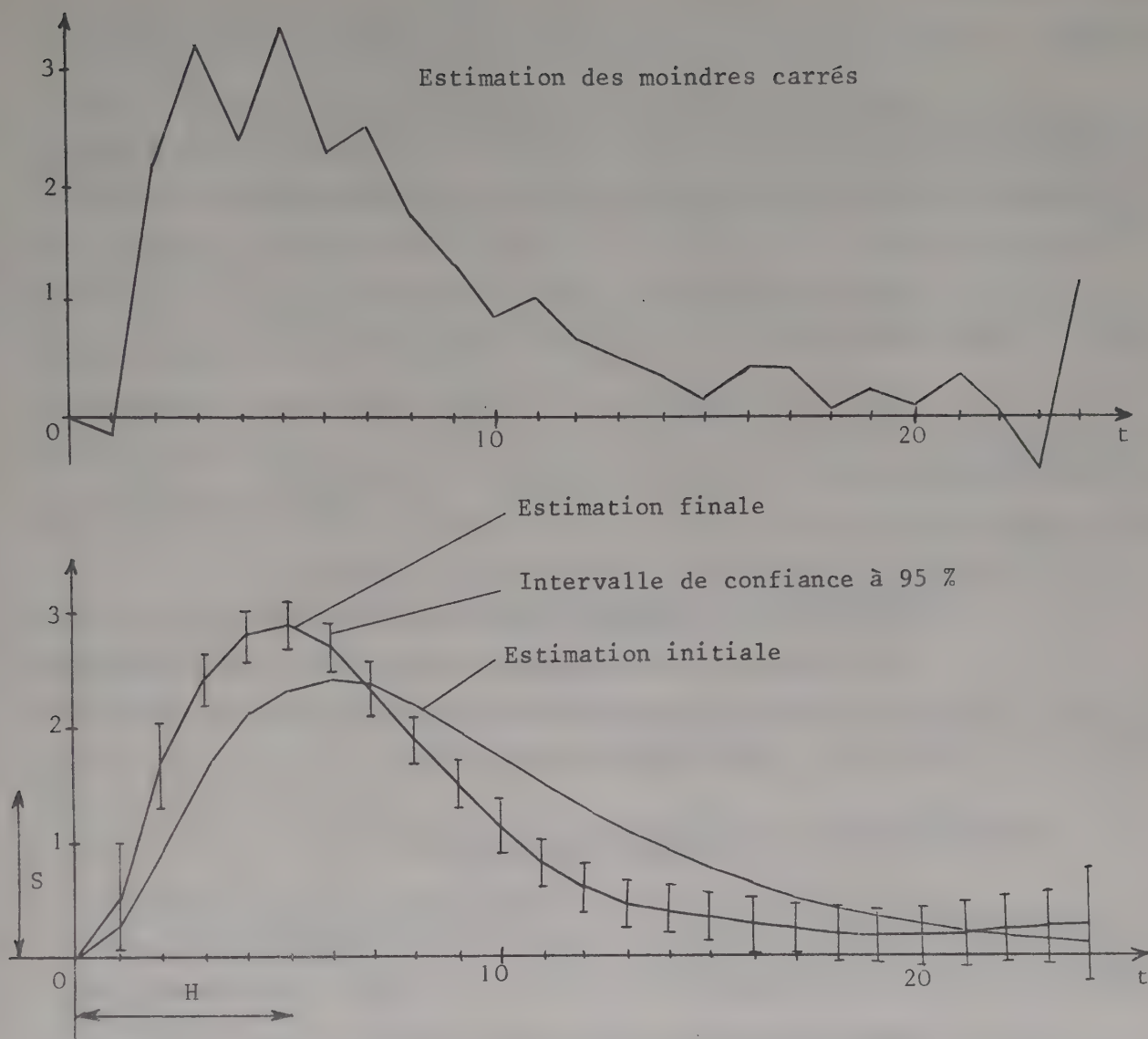


Fig. 6-7

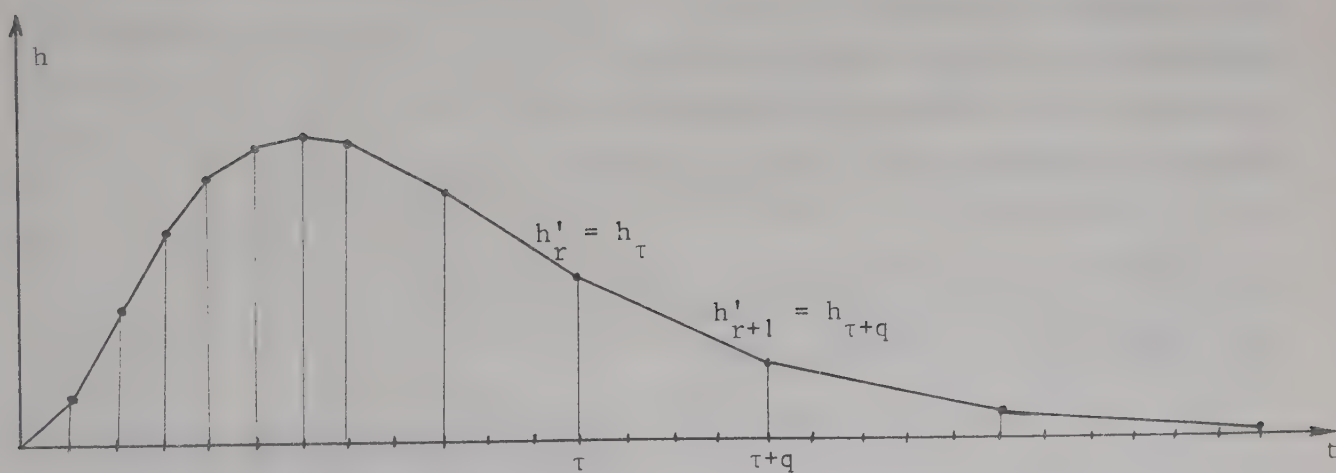


Fig. 6-8

Ce qui correspond à une estimation initiale volontairement éloignée de la courbe exacte.

La matrice de covariance initiale est $M_{ij} = e^{-\left(\frac{i-j}{5}\right)^2}$, soit un facteur de continuité plutôt faible, sans être négligeable, et un domaine de vraisemblance à 95,5 % pour $\underline{h}$ égal à $\hat{\underline{h}}_0 \pm 2$, ce qui englobe bien la courbe exacte.

On a hachuré sur la figure le domaine défini par l'estimation finale $\pm$ l'écart type déduit de la matrice de covariance finale.

Nous constatons que, si le début de la courbe est correctement identifié, les dernières valeurs restent proches de l'estimation initiale, mais cette estimation est accompagnée d'un écart type important qui montre le peu de signification de cette partie de l'estimation.

D) Identification d'un processus réel (fig. 6-7)

Nous reportons les résultats obtenus sur le désurchauffeur de l'unité pilote de l'E.N.S.M., en comparant avec l'identification obtenue en l'absence de connaissance à priori (moindres carrés).

6-1-4 Conclusions et extensions de la méthode

Nous avons remarqué dans le paragraphe 6-1-2 que l'on pouvait envisager le cas d'un bruit de mesure "coloré". Il est immédiat qu'on peut également identifier une composante continue sur la sortie (cf Barraud, 1970).

La méthode s'étend également au cas des systèmes multivariables (une sortie, plusieurs entrées).

Des obstacles matériels s'élèvent alors en raison de la dimension de la matrice de variance covariance (volume de mémoire et de calcul). Pour diminuer la dimension du problème, nous avons envisagé de considérer un échantillonnage à pas variable des réponses impulsionnelles, en interpolant linéairement les séquences $\underline{h}$ avec un pas de plus en plus large au fur et à mesure qu'on approche de la fin de la réponse (fig. 6-8). En identifiant seulement la sous-séquence $\underline{h}'$, on réduit ainsi dans un rapport allant jusqu'à 3 ou 4 la dimension du problème, ce qui permet de traiter des identifications multivariables (jusqu'à 5 entrées, 1 sortie) sur calculateur de moyenne importance (IBM 360-30, 64 K). cf Barraud, 1971.

Parmi les identifications effectuées, mentionnons l'identification multivariable de l'unité pilote du laboratoire d'automatique de l'E.N.S.M., un micro réseau de distribution (EDF), un processus chimique industriel (Collin, 1970), et la chaîne de tangage d'un avion (type Piaggio).

6-2 IDENTIFICATION PAR FONCTIONS MODULATRICES ET FILTRAGE LINEAIRE -

6-2-1 Rappel du principe de l'identification par fonctions modulatrices

Pour alléger les notations, nous considérerons un système du second ordre : les résultats s'étendent directement à un ordre quelconque.

Soient à identifier les quatre paramètres a_1 , a_2 , b_0 , b_1 d'un système représenté par l'équation différentielle :

$$x + a_1 \dot{x} + a_2 \ddot{x} = b_0 u + b_1 \dot{u} \quad 6-2-1$$

Les signaux d'entrée et de sortie : $u(t)$ et $x(t)$ sont mesurés expérimentalement depuis $t = 0$ à $t = T$.

On considère des fonctions $f_i(t)$ dites modulatrices dont les dérivées successives sont connues (à l'inverse de celles de u et x) et telles que $f_i(0)$, $\dot{f}_i(0)$, $f_i(T)$, $\dot{f}_i(T)$ soient nuls.

En multipliant les deux membres de (6-2-1) par $f_i(t)$ et en intégrant par parties, on démontre immédiatement que :

$$\int_0^T x f_i dt - a_1 \int_0^T x \dot{f}_i dt + a_2 \int_0^T x \ddot{f}_i dt = b_0 \int_0^T u f_i dt - b_1 \int_0^T u \dot{f}_i dt$$

Nous pouvons donc calculer les coefficients d'une équation :

$$y_{oi} + a_1 y_{1i} + a_2 y_{2i} = b_0 v_{oi} + b_1 v_{1i} \quad 6-2-2$$

En considérant quatre fonctions modulatrices distinctes, on peut obtenir ainsi un système d'équations permettant de calculer les quatre inconnues a_1 , a_2 , b_0 et b_1 .

La méthode peut être utilisée directement sous cette forme : elle demande alors de mettre en mémoire l'échantillon d'entrée-sortie, de faire le produit des fonctions x et u par les fonctions modulatrices f_i et leurs dérivées $\dot{f}_i$, $\ddot{f}_i$, et enfin d'effectuer les intégrations requises.

Remarquons que le calcul hybride est particulièrement indiqué pour ces opérations (Bonnemay, 1971).

D'autre part, il faut faire un choix parmi toutes les fonctions modulatrices possibles. A ce propos, nous savons qu'à partir d'une fonction $f(t)$ unique, nulle pour $t < 0$ et $t > T'$, avec $T' < T$, on peut obtenir différentes fonctions modulatrices (fig. 6-9).

Nous noterons :

$$f_i(t) = f(t_i - t) \quad (\text{où } T' < t_i < T)$$

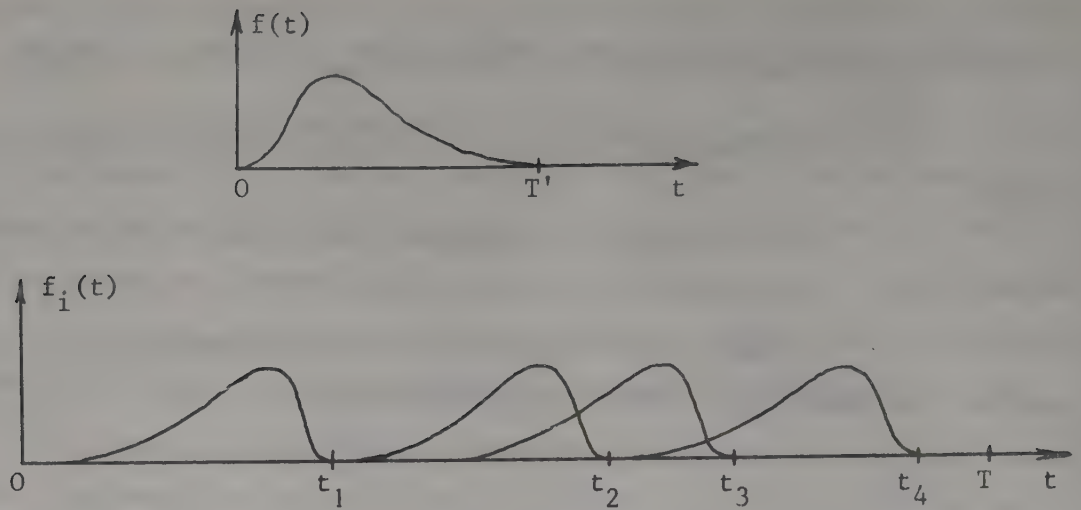


Fig. 6-9

6-2-2 Nouvelle formulation des résultats

Notons $(f * g)$ le produit de convolution : $\int_{-\infty}^{+\infty} g(\tau) f(t-\tau) d\tau$

Définissons : $y_0(t) = (x * f)$, $y_1(t) = (x * \dot{f})$, ... , $v_0(t) = (u * f)$, ...

Alors : $y_{0i} = y_0(t_i)$, $y_{1i} = y_1(t_i)$, ... , $v_{0i} = v_0(t_i)$, ...

Nous pouvons donc obtenir, au besoin *en temps réel*, les coefficients y_{0i} , y_{1i} , y_{2i} , v_{0i} , v_{2i} de l'équation (6-2-1).

Il suffit de traiter $x(t)$ et $u(t)$ par des filtres ayant pour réponse impulsionnelle $f(t)$, $\dot{f}(t)$, $\ddot{f}(t)$: (fig. 6-10).

On évite ainsi tout l'ensemble des opérations de mise en mémoire, de produits de fonction et d'intégration.

D'autre part, on peut remarquer que $v_1 = \dot{v}_0$, $y_1 = \dot{y}_0$, $y_2 = \ddot{y}_0$. Ceci peut simplifier la réalisation des filtres : considérons en effet le schéma analogique de la fig. 6-11. Nous voyons qu'on peut obtenir y_0 , y_1 , y_2 à partir d'un filtre unique. Il suffit donc de *deux* filtres, un pour l'entrée et un pour la sortie, pour obtenir l'ensemble des coefficients y et v .

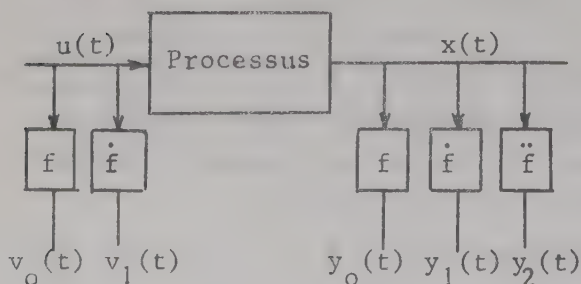


Fig. 6-10

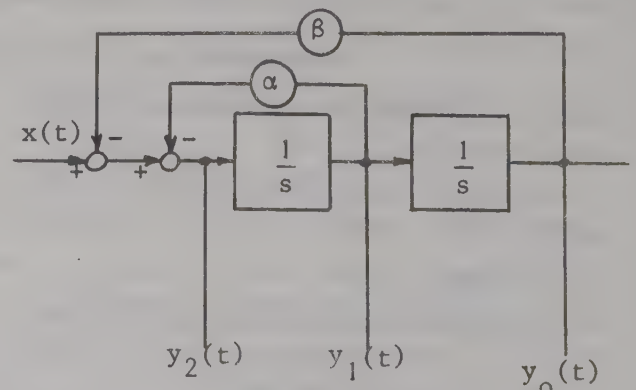


Fig. 6-11

6-2-3 Traitement des données filtrées

Nous disposons finalement d'une équation à coefficients variables connus : $y_0(t) = -y_1(t) a_1 - y_2(t) a_2 + v_0(t) b_0 + v_1(t) b_1$

On peut alors prélever à des instants discrets ($t=t_1, t_2, \dots, t_i, \dots$) les coefficients $y_0(t)$, $y_1(t)$ de l'équation 6-2-3 et traiter le système surabondant obtenu par une méthode de moindres carrés. Il apparaît d'ailleurs tout indiqué d'utiliser les moindres carrés sous leur forme récurrente, ce qui permet alors une identification en temps réel. Notons aussi qu'on peut raffiner l'algorithme afin de détecter d'éventuelles variations de paramètres en suivant un mode opératoire analogue à celui qui est proposé par Aguilar (1970).

Nous pouvons également envisager un algorithme continu.

Notons $C(t) = [-y_1(t), -y_2(t), v_0(t), v_1(t)]$

$$\underline{\theta} = [a_1, a_2, b_0, b_1]^T$$

Alors $y_0 = C \underline{\theta}$

L'algorithme le plus simple qu'on puisse proposer est le suivant :

$$\dot{\hat{\theta}} = C^T k (y_0 - C \hat{\theta})$$

où k est un facteur positif arbitraire.

On démontre immédiatement que l'algorithme est stable.

En effet, soit $\underline{\Delta} = \underline{\theta} - \hat{\underline{\theta}}$

Alors $\dot{\underline{\Delta}} = -C^T k C \underline{\Delta}$

$C^T k C$ étant une matrice définie non négative, la somme des carrés des écarts $(\underline{\theta} - \hat{\underline{\theta}})$ est non croissante et l'algorithme est stable.

Nous ne disposons pas d'une démonstration de la stabilité asymptotique ($\underline{\Delta} \rightarrow 0$ quand $t \rightarrow \infty$), cependant il semble que celle-ci soit assurée sous des conditions très générales et réalisées par tout signal d'entrée non trivial (autre qu'une constante ou un signal harmonique pur par exemple).

6-2-4 Choix du filtre

Ce problème devrait être traité en parallèle avec l'étude de l'influence des bruits de mesure. Disons seulement que, par analogie avec l'identification des systèmes discrets par la méthode de Steiglitz (1965), on peut proposer de choisir, pour un système à identifier de transmittance $N(s)/D(s)$, un filtre f dont la transmittance soit voisine de $1/D(s)$.

6-2-5 Résultats expérimentaux

Nous avons simulé sur calculateur analogique (Analac A 110) un système du second ordre :

$$H(s) = \frac{b_0}{1 + a_1 s + a_2 s^2}$$

$$\begin{aligned} b_0 &= 1 \\ a_1 &= 1 \\ a_2 &= 2 \end{aligned}$$

Le signal d'excitation $u(t)$ est fourni par un générateur de bruit, d'où un signal de sortie $x(t)$. Les trois coefficients b_0 , a_1 , a_2 sont identifiés en temps réel à partir du couple $u(t)$, $x(t)$.

Nous avons étudié le comportement de l'identification dans des conditions variées et nous présentons ici quelques uns des résultats obtenus, concernant essentiellement l'influence du gain sur l'algorithme d'identification :

$$\dot{\hat{\theta}} = C^T k (y_0 - C \hat{\theta})$$

- figure 6-12 : résultats d'une identification pour une valeur type k_0 .

- figure 6-13 : résultats avec $k = k_0/10$. On constate alors que l'algorithme converge beaucoup plus lentement (noter le changement de l'échelle des temps sur l'enregistrement).

- figure 6-14 : le gain $k = k_0 \times 10$ conduisant à la saturation de certains organes du calculateur, nous nous sommes limités à $k = k_0 \times 3$.

Nous constatons que la rapidité de convergence est approximativement proportionnelle à K , mais que l'écart type de l'erreur finale croît également proportionnellement à k . L'expérimentation permet cependant facilement de trouver la valeur k_0 conduisant à un compromis acceptable entre rapidité et précision.

On peut évidemment envisager de prendre un gain k variable au cours du temps, pour concilier rapidité et précision. Une décroissance de k en $\frac{1}{t}$, satisfaisant aux conditions générales de l'approximation stochastique (Albert, 1967) semble indiquée. Cependant, s'il s'agit de systèmes non stationnaires, il est clair que le gain k ne doit pas tendre finalement vers 0 pour que l'algorithme puisse réagir à tout instant à des variations des coefficients.

Nous montrons sur la figure (6-15), pour $k = k_0$, le comportement des estimations lors d'une variation du paramètre a_2 .

Les essais présentés ici ont été réalisés en filtrant u et x par des fonctions de transfert identiques au dénominateur de la transmittance à identifier. D'autres essais ont été faits en modifiant ces filtres, en

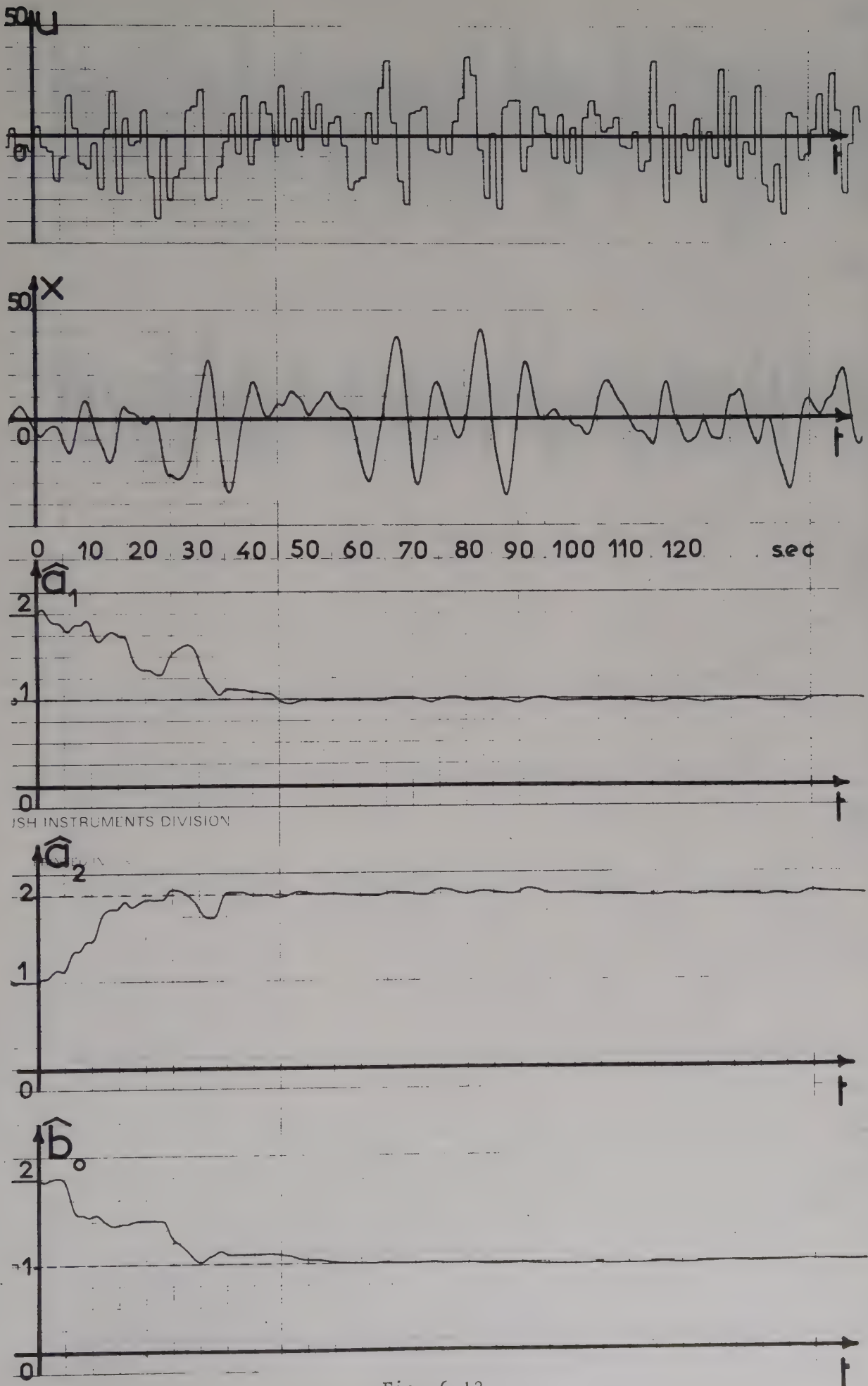


Fig. 6-12

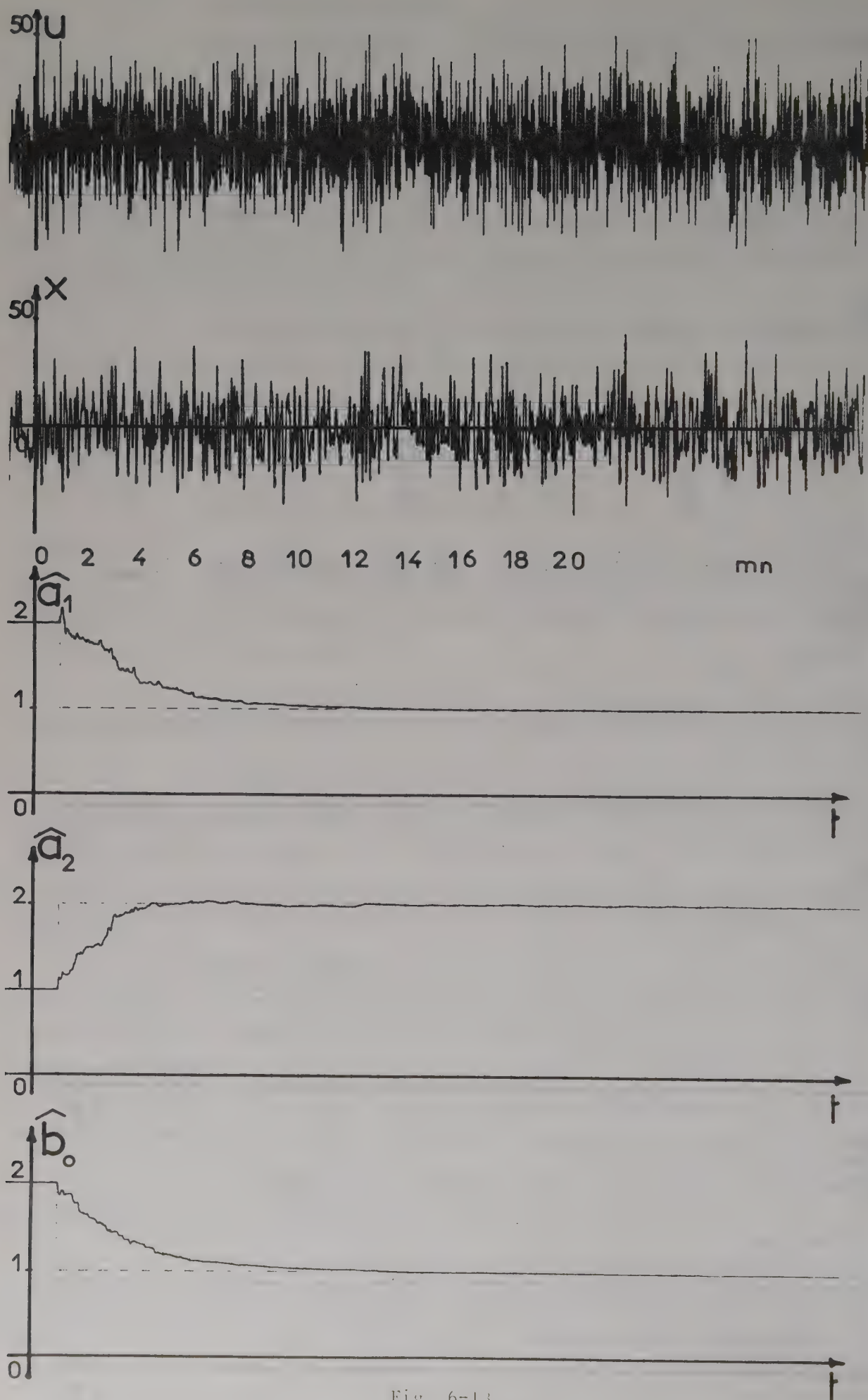


Fig. 6-13

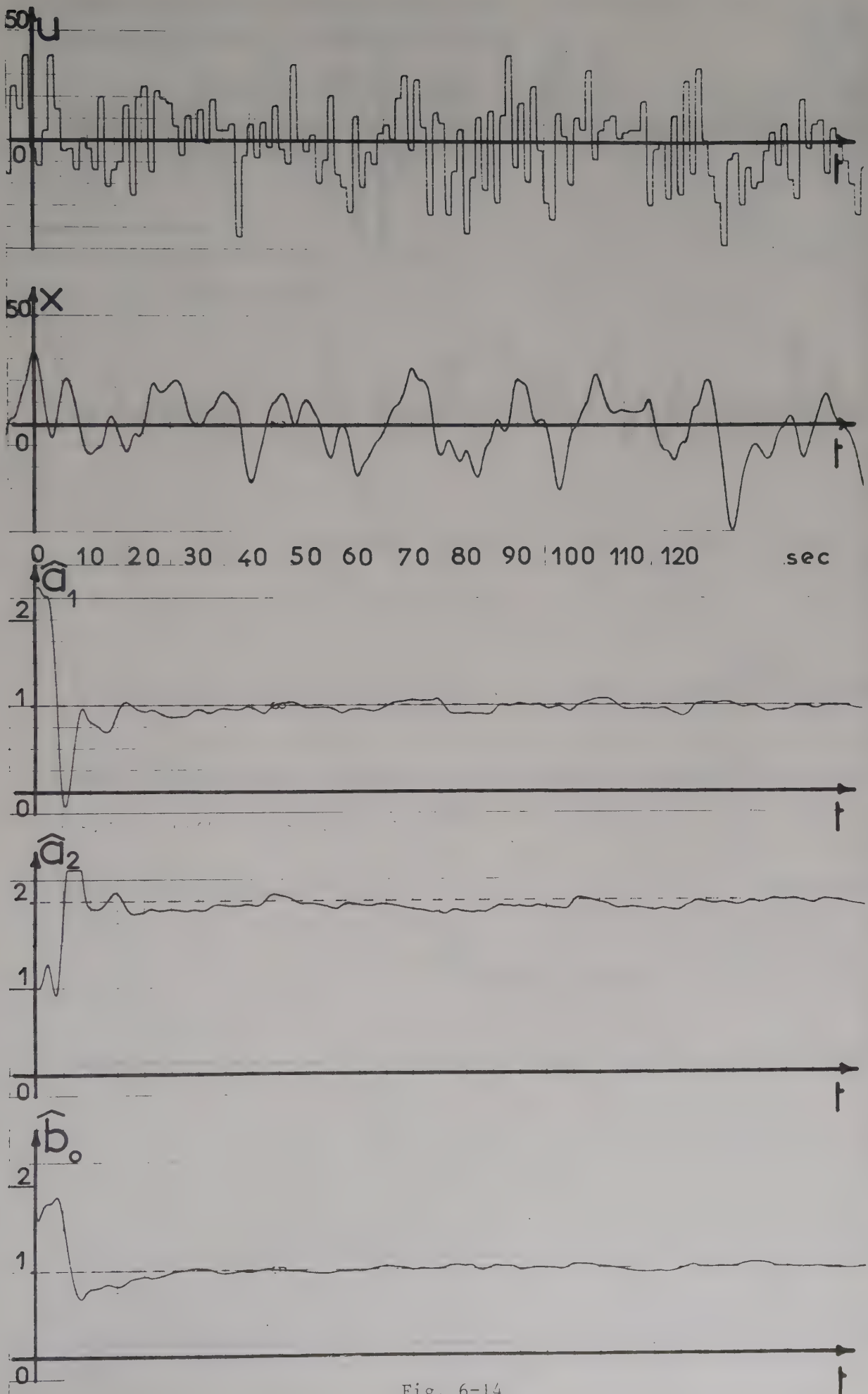


Fig. 6-14

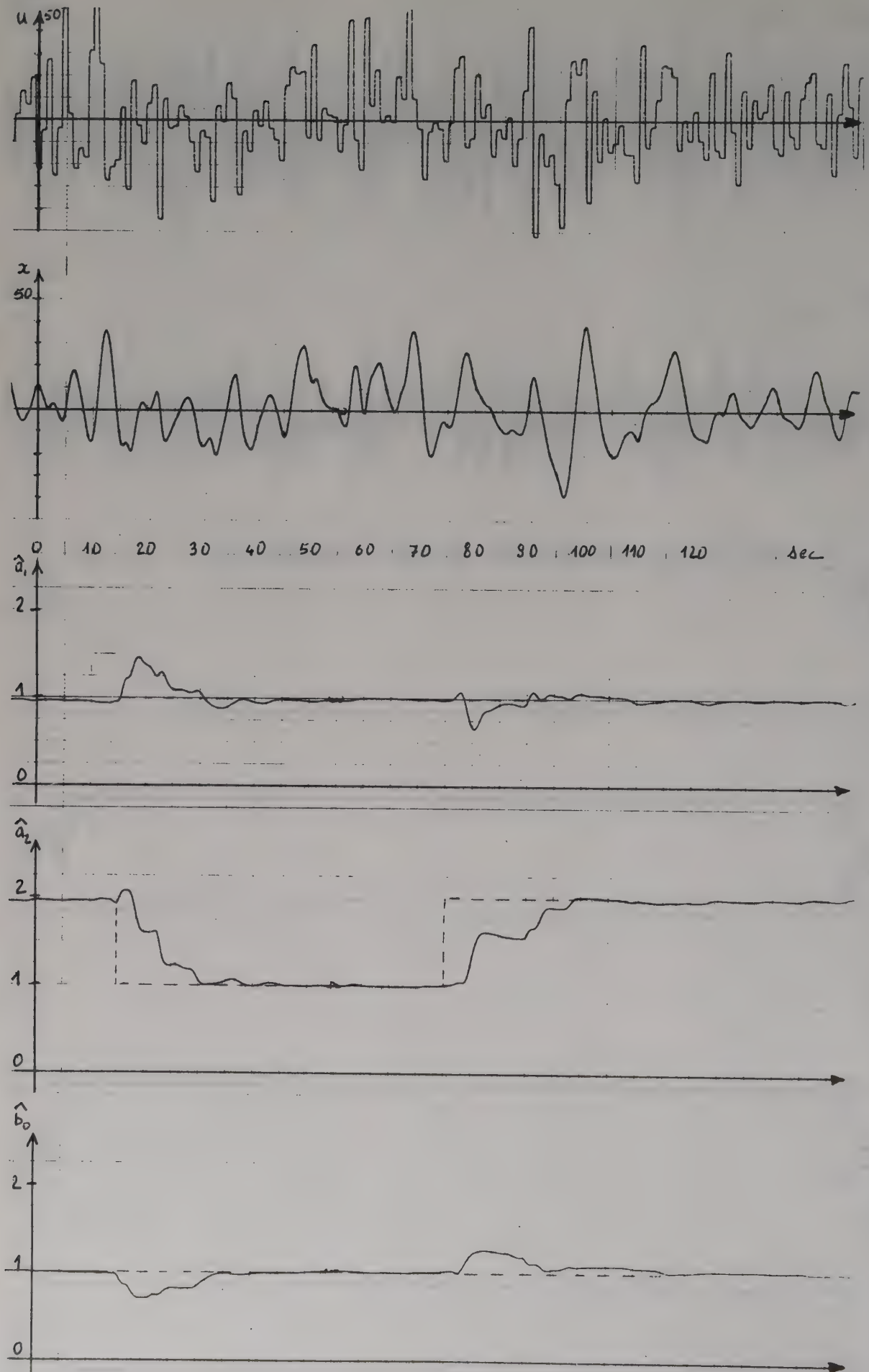


Fig. 6-15

choisissant des entrées différentes (on constate en particulier que l'algorithme ne converge pas pour une entrée sinusoïdale).

L'influence de ces différents choix, ainsi que les effets des bruits, restent évidemment à approfondir, aussi bien sur le plan théorique que sur le plan expérimental. Dans son état actuel, la méthode paraît néanmoins applicable et fournir des résultats intéressants puisque, pour un système ayant un temps de réponse d'une dizaine de secondes, des résultats convenables sont obtenus en quelques 30 secondes, en partant de valeurs initiales éloignées de 50 à 100 % des valeurs exactes.

C O N C L U S I O N

Le modèle dynamique le plus largement utilisé en automatique est certainement celui des équations différentielles linéaires à constantes localisées ou, sous forme discrète, des équations récurrentes linéaires. On conçoit donc que la plus grande partie des publications sur l'identification concerne ce type de modèle. Il résulte de ces travaux qu'à moins de considérer le problème d'un point de vue déterministe, comme nous l'avons fait dans le dernier chapitre (§ 6-2) ou d'envisager des cas particuliers comme celui des séquences de pondération (§ 6-1), l'identification se présente en général comme un problème non linéaire, pour lequel il n'existe pas de solution optimale par algorithme fini (§ 1-1).

A fortiori, en est-il de même pour l'identification des systèmes non linéaires. Devant cette situation, il n'est pas question de trancher catégoriquement entre les diverses approximations qui peuvent être proposées et qui possèdent chacune leurs avantages, selon qu'on considère telle ou telle classe de processus, selon aussi les besoins de l'utilisateur et les contraintes de calcul. Cependant, il nous semble que l'identification par filtrage non linéaire doit occuper une place convenable parmi les différentes théories actuelles.

Nous avons vu qu'un des attrait du filtrage non linéaire est d'abord de poser le problème de l'identification d'une manière synthétique, recouvrant le cas des systèmes non linéaires ou non stationnaires, à condition que l'on puisse représenter à la fois le processus proprement dit, l'évolution de ses paramètres, les perturbations et les bruits de mesure, par un modèle général sous forme d'équations différentielles ou récurrentes stochastiques (§ 2-2).

Il ne s'agit pas là d'une synthèse destinée simplement à quelque satisfaction de l'esprit, mais d'un point de départ permettant de parvenir à des procédés de calculs effectifs, éventuellement applicables en temps réel.

Ces procédés se sont avérés fonctionner, non seulement sur de simples simulations illustratives, mais sur des cas concrets et variés (processus chimiques, thermiques, aéronautiques, etc...), recouvrant bien entendu les indispensables modèles linéaires (mono et multivariables, continus ou discrets), mais aussi des systèmes non linéaires (§ 5-4). Toutes ces identifications ont d'ailleurs été effectuées en adaptant à chaque cas particulier un petit nombre

de programmes généraux (§ 3-1) qui permettent de mettre en oeuvre le filtrage non linéaire avec un minimum de travail d'analyse. Quant au temps de calcul, sur le seul exemple où nous ayons pu faire la comparaison (turbo réacteur), il s'est montré six fois inférieur à celui de la méthode du modèle.

Notons, cependant, quelques limitations au champ des applications. Les non linéarités de l'équation d'état généralisée introduite pour l'identification doivent être deux fois différentiables (§ 2-1), ce qui exclut par exemple l'identification des caractéristiques de systèmes "tout ou rien". D'autre part, les bruits et perturbations sont supposés gaussiens, hypothèse souvent vérifiée en pratique, au moins approximativement, et elle n'est pas impérative pour l'interprétation déterministe de l'algorithme comme formulation récurrente approchée de la méthode du modèle (§ 2-3).

La restriction la plus sévère se situe sans doute au niveau de la dimension du problème. "Sans doute" car personnellement nous n'avons pas rencontré cette limitation : nous avons identifié par exemple les 7 paramètres d'un système non linéaire sans rencontrer de difficultés particulières (§ 5-4). Cependant, il ne peut y avoir de miracle : les difficultés de programmation non linéaire qu'on observe par exemple en usant de la méthode du modèle sont peut-être repoussées, mais doivent nécessairement se retrouver, sous une forme ou sous une autre, dans le filtrage non linéaire.

Nous ne pensons pas qu'il soit possible d'identifier simultanément beaucoup plus d'une dizaine de paramètres (pour un processus dynamique unitaire évidemment : si on peut le décomposer en plusieurs sous-systèmes de dimension réduite, il n'y a plus de limite).

Pour éviter la nécessité de déterminer exactement la variance des bruits et perturbations, nécessaires à l'établissement de l'algorithme, nous avons présenté (chapitre 4) une méthode d'estimation de celles-ci, applicable simultanément à l'identification des paramètres. En dehors d'une simulation (§ 4-4), nous n'avons pratiquement testé cette méthode que dans le cas de l'estimation de bruit de mesure (§ 5-1). Dans le cas où s'ajoutent des perturbations internes, son application semble plus délicate, non seulement du point de vue pratique, mais encore sur le plan théorique où il apparaît nécessaire d'approfondir l'identifiabilité par filtrage non linéaire approché des variances covariances de perturbations multivariées.

Le principal point de discussion reste certainement celui de l'initialisation et de la convergence de l'algorithme.

D'une manière générale, la convergence des filtres non linéaires approximativement optimaux n'est jamais assurée, mais la probabilité de diver-

gence est d'autant plus faible que l'estimation initiale est meilleure. La nécessité d'une initialisation soulève d'abord la difficulté de disposer de celle-ci (connaissances à priori) et ensuite l'objection que les estimations finales dépendent des valeurs initiales choisies. Nous répondrons alors que dans presque tous les cas que nous avons traités, nous avons constaté que la méthode convergeait malgré des initialisations suffisamment "larges" pour que le résultat final soit pratiquement indépendant de l'initialisation : nous ne nous sommes jamais trouvés devant un problème où à la fois les connaissances à priori seraient complètement absentes, et où quelques essais ne permettraient pas de trouver un point de départ convenable pour effectuer l'identification.

Le fait que la convergence ne soit pas certaine serait plus inquiétant lors d'une identification en temps réel : soit que de précieuses mesures, en l'absence d'enregistrement, soient perdues à la suite d'un mauvais fonctionnement du filtre, soit que des résultats d'identification aberrants, utilisés en ligne pour l'adaptation d'un procédé de commande, puissent provoquer quelque catastrophe. Ces craintes sont cependant peu fondées dans la mesure où il n'est pas concevable d'aborder une identification en temps réel sans quelque connaissance préalable du processus, et notre expérience nous permet d'avancer que pour une initialisation raisonnable de l'algorithme, les probabilités de divergence sont réellement très faibles.

Quant à l'influence de l'initialisation sur le résultat final, nous avons mis au point, dans une optique "hors ligne", des modalités d'élimination des connaissances à priori (§ 3-3) : cette élimination ne laisse subsister dans le résultat final que la contribution des informations apportées par les mesures. Ceci permet alors d'initialiser au mieux l'algorithme et d'en faciliter ainsi la convergence, sans que cette initialisation soit prise en compte dans l'estimation finale des paramètres.

Nous concluons en disant que s'il est nécessaire de prendre quelques soins lors de l'initialisation du filtre non linéaire, on se trouve largement récompensé de cet effort par la puissance et la souplesse de cette méthode d'identification.

LE FILTRE DE KALMAN DISCRET

Il existe de multiples points de vue et de multiples façons d'établir les équations du filtre de Kalman (cf Radix, 1970).

Dans l'optique de notre travail, nous allons présenter celles-ci vues sous l'angle des probabilités conditionnelles.

- Position du problème

Soit un système linéaire récurrent stochastique :

$$\underline{x}_{t+1} = A \underline{x}_t + B \underline{u}_t + \underline{\delta}_t$$

$$\underline{s}_{t+1} = C_{t+1} \underline{x}_{t+1} + \underline{\delta}'_{t+1}$$

$\underline{\delta}_t$ et $\underline{\delta}'_t$ sont des séquences indépendantes gaussiennes centrées :

$$E(\underline{\delta}_t / \underline{x}_t, \underline{x}_{t-1}, \dots) = \underline{0}$$

$$E(\underline{\delta}'_t / \underline{x}_t, \underline{x}_{t-1}, \dots) = \underline{0}$$

$$E(\underline{\delta}_t \underline{\delta}_{t+\tau}^T) = \begin{cases} 0 & \text{pour } \tau \neq 0 \\ \Phi_t & \text{pour } \tau = 0 \end{cases}$$

$$E(\underline{\delta}'_t \underline{\delta}'_{t+\tau}^T) = \begin{cases} 0 & \text{pour } \tau \neq 0 \\ \Phi'_t & \text{pour } \tau = 0 \end{cases}$$

$$E(\underline{\delta}'_t \underline{\delta}_{t+\tau}) = 0 \quad \forall \tau$$

Le problème consiste à déterminer :

$$E(\underline{x}_t / \dots, \underline{s}_{t-1}, \underline{s}_t), \text{ noté } E(\underline{x}_t / t) = \hat{\underline{x}}_t$$

On passera par la grandeur intermédiaire :

$$E(\underline{x}_t / \dots, \underline{s}_{t-1}), \text{ noté } E(\underline{x}_t / t-1) = \underline{x}_t^*$$

On note encore :

$$\hat{M}_t = \text{Var} (\underline{x}_t / t) = E[(\underline{x}_t - \hat{\underline{x}}_t)(\underline{x}_t - \hat{\underline{x}}_t)^T / t]$$

$$M_t^* = \text{Var} (\underline{x}_t / t-1) = E[(\underline{x}_t - \underline{x}_t^*)(\underline{x}_t - \underline{x}_t^*)^T / t-1]$$

Pour établir les équations récurrentes permettant de calculer $\hat{\underline{x}}$, $\hat{\underline{M}}$, $\underline{x}^*$ et $\underline{M}^*$, nous allons utiliser deux résultats intermédiaires.

———A) Soit $\underline{y}$ un vecteur gaussien d'espérance $\underline{y}' = E(\underline{y})$ et de variance $\underline{M}'_{yy} = E[(\underline{y}-\underline{y}') (\underline{y}-\underline{y}')^T]$.

Considérons la variable $\underline{z} = A \underline{y} + \underline{b}$ (A et $\underline{b}$ non aléatoires). On sait alors que $\underline{z}$ est gaussien.

Avec

$$\begin{aligned} E(A \underline{y} + \underline{b}) &= E(\underline{z}) = \underline{z}' = A \underline{y}' + \underline{b} \\ E[A(\underline{y}-\underline{y}') (\underline{y}-\underline{y}')^T A^T] &= E[(\underline{z}-\underline{z}') (\underline{z}-\underline{z}')^T] = \underline{M}'_{zz} = A \underline{M}'_{yy} A^T \end{aligned}$$

———B) Soit $\underline{y}$ et $\underline{z}$ deux vecteurs gaussiens d'espérance $\underline{y}'$ et $\underline{z}'$, de variances et de covariance $\underline{M}'_{yy}$, $\underline{M}'_{zz}$, et $\underline{M}'_{yz} = E[(\underline{y}-\underline{y}') (\underline{z}-\underline{z}')^T]$.

On peut écrire les densités de probabilité

$$p(\underline{y}, \underline{z}) = k_1 \exp \left(-\frac{1}{2} \begin{bmatrix} \underline{y}-\underline{y}' \\ \underline{z}-\underline{z}' \end{bmatrix}^T \underline{M}'^{-1} \begin{bmatrix} \underline{y}-\underline{y}' \\ \underline{z}-\underline{z}' \end{bmatrix} \right)$$

avec :

$$\underline{M}' = \begin{bmatrix} \underline{M}'_{yy} & \underline{M}'_{yz} \\ \underline{M}'_{yz} & \underline{M}'_{zz} \end{bmatrix}$$

et :

$$k_1 = 1/\sqrt{2}^n \cdot \sqrt{\text{Det}(\underline{M}')}^n$$

$$p(\underline{z}) = k_2 \exp \left(-\frac{1}{2} (\underline{z}-\underline{z}')^T \underline{M}'_{zz}^{-1} (\underline{z}-\underline{z}') \right)$$

Alors l'axiome des probabilités conditionnelles s'écrit :

$$p(\underline{y}/\underline{z}) = \frac{p(\underline{y}, \underline{z})}{p(\underline{z})} = k_3 \exp \left(-\frac{1}{2} \begin{bmatrix} \underline{y}-\underline{y}' \\ \underline{z}-\underline{z}' \end{bmatrix}^T \underline{M}'^{-1} \begin{bmatrix} \underline{y}-\underline{y}' \\ \underline{z}-\underline{z}' \end{bmatrix} + \frac{1}{2} (\underline{z}-\underline{z}')^T \underline{M}'_{zz}^{-1} (\underline{z}-\underline{z}') \right)$$

L'argument de l'exponentielle étant encore une fonction quadratique de $\underline{y}$, la variable $(\underline{y}/\underline{z})$ est encore gaussienne et nous allons montrer que :

$$p(\underline{y}/\underline{z}) = k_4 \exp \left(-\frac{1}{2} (\underline{y}-\underline{y}'')^T \underline{M}''_{yy}^{-1} (\underline{y}-\underline{y}'') \right)$$

avec :

$$\begin{aligned} \underline{y}'' &= \underline{y}' + \underline{M}'_{yz} \underline{M}'_{zz}^{-1} (\underline{z}-\underline{z}') = E(\underline{y}/\underline{z}) \\ \underline{M}''_{yy} &= \underline{M}'_{yy} - \underline{M}'_{yz} \underline{M}'_{zz} \underline{M}'_{zy} = \text{Var}(\underline{y}/\underline{z}) \end{aligned}$$

Démonstration :

Notons $\underline{M}'^{-1} = \begin{bmatrix} U & W \\ W^T & V \end{bmatrix}$

($\underline{M}'$ étant symétrique, $\underline{M}'^{-1}$ est symétrique)

soit q l'argument de l'exponentielle de $p(\underline{y}/\underline{z})$

$$q = -\frac{1}{2} [(\underline{y}-\underline{y}')^T U(\underline{y}-\underline{y}') + 2(\underline{y}-\underline{y}')^T W(\underline{z}-\underline{z}') + (\underline{z}-\underline{z}')^T (V-M'_{zz}^{-1}) (\underline{z}-\underline{z}')]]$$

qu'il s'agit d'identifier avec :

$$q = -\frac{1}{2} (\underline{y}-\underline{y}'')^T M''_{yy}^{-1} (\underline{y}-\underline{y}'')$$

Dans la deuxième expression, q est minimum pour $\underline{y} = \underline{y}''$. Dérivons la première par rapport à $\underline{y}$:

$$\underline{q}_y = -U(\underline{y}-\underline{y}') - W(\underline{z}-\underline{z}')$$

$$\underline{q}_y = \underline{0} \text{ pour } \underline{y}-\underline{y}' = -U^{-1} W(\underline{z}-\underline{z}')$$

de l'égalité $M'^{-1} M' = 1$, on tire :

$$U M'_{yz} + W M'_{zz} = 0$$

$$\text{d'où : } U^{-1} W = -M'_{yz} M'^{-1}_{zz}$$

$$\text{d'où } \underline{q}_y = \underline{0} \text{ pour } \underline{y} = \underline{y}'' = \underline{y}' + M'_{yz} M'^{-1}_{zz} (\underline{z}-\underline{z}')$$

De même, nous établissons l'égalité de la matrice des dérivées secondes, par rapport à $\underline{y}$, des deux expressions de q

de la première expression, on tire :

$$Q_{yy} = -U$$

de la deuxième :

$$Q_{yy} = -M''_{yy}^{-1}$$

$$\text{d'où : } M''_{yy} = U^{-1}$$

de l'égalité $M'^{-1} M = 1$, on tire :

$$U M'_{yy} + W M'_{yz} = 1$$

$$\text{d'où : } M'_{yy} + U^{-1} W M'_{yz} = U^{-1}$$

$$\text{en reprenant } U^{-1} W = -M'_{yz} M'^{-1}_{zz}$$

$$M'_{yy} - M'_{yz} M'^{-1}_{zz} M'_{zy} = U^{-1}$$

$$\text{on trouve donc bien : } M''_{yy} = M'_{yy} - M'_{yz} M'^{-1}_{zz} M'_{zy}$$

Les résultats intermédiaires ci-dessus permettent d'établir immédiatement les équations du *filtre de Kalman*.

C) Reprenons l'équation :

$$\underline{x}_{t+1} = A_t \underline{x}_t + B_t \underline{u}_t + \underline{\delta}_t$$

Notons $\underline{y}$ le vecteur aléatoire $\begin{bmatrix} \underline{x}_t \\ \underline{\delta}_t \end{bmatrix}$ conditionnellement à t

$$\underline{y}' = \begin{bmatrix} \hat{\underline{x}}_t \\ 0 \end{bmatrix}, M'_{yy} = \begin{bmatrix} \hat{M}_t & 0 \\ 0 & \Phi_t \end{bmatrix}$$

Notons $\underline{z}$ le vecteur aléatoire $\underline{x}_{t+1}$ conditionnellement à t

$$\underline{z} = \begin{bmatrix} A_t & 1 \end{bmatrix} \underline{y} + \underline{b} \quad \text{avec } \underline{b} = B_t \underline{u}_t$$

Alors, d'après le résultat intermédiaire (A)

$$\begin{aligned} \underline{z}' &= E(\underline{x}_{t+1}/t) = \underline{x}_{t+1}^* = A_t \hat{\underline{x}}_t + B_t \underline{u}_t \\ M'_{zz} &= \text{Var}(\underline{x}_{t+1}/t) = M_{t+1}^* = A_t \hat{M}_t A_t^T + \Phi_t \end{aligned}$$

ce qui démontre les équations de "prédiction".

D) Reprenons l'équation :

$$\underline{s}_{t+1} = C_{t+1} \underline{x}_{t+1} + \underline{\delta}'_{t+1}$$

Notons maintenant $\underline{y}$ le vecteur aléatoire $\underline{x}_{t+1}$ conditionnellement à t

$\underline{z}$ le vecteur aléatoire $\underline{s}_{t+1}$ conditionnellement à t

$$\text{alors : } \underline{y}' = \underline{x}_{t+1}^*, M'_{yy} = M_{t+1}^*$$

et en utilisant le résultat intermédiaire (A) :

$$\underline{z}' = C_{t+1} \underline{x}_{t+1}^*, M'_{zz} = C_{t+1} M_{t+1}^* C_{t+1}^T + \Phi'_{t+1}$$

$$\text{de plus : } M'_{yz} = M_{t+1}^* C_{t+1}^T$$

en utilisant alors le résultat intermédiaire (B) :

$$\begin{aligned} \underline{y}'' &= E(\underline{y}/\underline{z}) = E(\underline{x}_{t+1}/t+1) = \hat{\underline{x}}_{t+1} = \underline{x}_{t+1}^* + M_{t+1}^* C_{t+1}^T (C_{t+1} M_{t+1}^* C_{t+1}^T + \Phi'_{t+1})^{-1} (\underline{s}_t - C_{t+1} \underline{x}_{t+1}^*) \\ M''_{yy} &= \text{Var}(\underline{y}/\underline{z}) = \text{Var}(\underline{x}_{t+1}/t+1) = \hat{M}_{t+1} = M_{t+1}^* - M_{t+1}^* C_{t+1}^T (C_{t+1} M_{t+1}^* C_{t+1}^T + \Phi'_{t+1})^{-1} C_{t+1} M_{t+1}^* \end{aligned}$$

ce qui démontre les équations "d'estimation".

Remarques :

- $(\underline{x}_{t+1}/t+1)$ est une variable gaussienne si $(\underline{x}_t/t)$ est lui-même une variable gaussienne. Il est donc nécessaire que l'estimation initiale $(\underline{x}_0/0)$ soit gaussienne.

- La démonstration (B) que nous avons faite suppose que la matrice M' est régulière. Les résultats subsistent néanmoins même si M' est singulière.

Annexe II

DESCRIPTION ET LISTING DES SOUS-PROGRAMMES D'IDENTIFICATION

Nous considérons un vecteur aléatoire de dimension N, supposé gaussien, défini par son vecteur valeur moyenne Z et sa matrice de covariance C.

L'objet de ces sous-programmes est de remplacer Z et C, qui seront à la fois arguments d'entrée et de sortie, par la valeur moyenne et la covariance d'un nouveau vecteur aléatoire qui sera :

- soit une fonction de l'ancien vecteur,
- soit l'ancien vecteur conditionnellement à une mesure.

Dans le texte qui suit, la notation Z désignera indifféremment le vecteur aléatoire ou sa valeur moyenne, étant entendu que dans les programmes proprement dits, le tableau Z est occupé par l'espérance de Z.

I - SUBROUTINE CADDE (C,Z,M,N,FI,I,FJ,J)

Arguments d'entrée : C,Z,M,N,FI,I,FJ,J

Arguments de sortie : C,Z

Objet :

Remplace la $I^{i\text{ème}}$ composante du vecteur aléatoire Z par une nouvelle variable, égale à $FI.Z(I) + FI.Z(J)$.

Précisons bien qu'on calcule non seulement l'espérance et la variance de la nouvelle composante, mais sa covariance avec les autres composantes qui, elles, demeurent inchangées. En d'autres termes :

$$Z = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & \ddots & & \\ 0 & \dots & FJ & \dots & FI & \dots & 0 \\ & & & \ddots & & & \\ & & & & \ddots & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} Z$$

$\begin{array}{c} \vdots \\ \vdots \\ \vdots \end{array} \begin{array}{l} \rightarrow I^{i\text{ème}} \text{ colonne} \\ \rightarrow J^{i\text{ème}} \text{ colonne} \end{array}$

(le signe = étant pris au sens du Fortran)

Remarques :

Le nouveau vecteur est une fonction linéaire de l'ancien, sa moyenne Z et sa covariance C sont donc données directement par le calcul intermédiaire A (cf. Annexe I).

L'argument d'entrée N est la dimension du vecteur Z et de la matrice C.

Dans certaines applications (cf. Annexe III), il est suffisant d'opérer seulement sur les M premières composantes de Z et C. Si $M < N$, alors $Z(K)$ et $C(K,L)$ demeurent alors inchangés pour K ou $L > M$. Cette remarque s'applique à tous les autres sous-programmes.

Listing :

```

SUBROUTINE CADDE(C,Z,M,N,FI,I,FJ,J)
C
  DIMENSION C(1),Z(1)
  Z(I)=FI*Z(I)+FJ*Z(J)
  II=N*(I-1)+I
  IJ=N*(J-1)+I
  JJ=N*(J-1)+J
  CC=FI*FI*C(II)+2*FI*FJ*C(IJ)+FJ*FJ
  1*C(JJ)
  IK=I-N
  JK=J-N
  NI=N*(I-1)
  DO 10 K=1,M
    IK=IK+N
    JK=JK+N
    KI=NI+K
    C(IK)=FI*C(IK)+FJ*C(JK)
  10 C(KI)=C(IK)
  C(II)=CC
  RETURN
  END

```

II - SUBROUTINE CPROD (C,Z,M,N,FI,I,FKL,F,L)

Arguments d'entrée : C,Z,M,N,FI,I,FKL,K,L

Arguments de sortie : C,Z

Objet :

Remplace la $I^{i\text{ème}}$ composante de Z par une nouvelle variable égale à $FI.Z(I) + FKL.Z(K).Z(L)$.

Le nouveau vecteur aléatoire Z est une fonction non linéaire de l'ancien vecteur, et nous développons alors l'équation de prédiction du filtrage non linéaire approché (§ 2-1) appliqué à ce cas particulier.

Remarque :

Sous-programme appelé : CADDE

Listing :

```

SUBROUTINE CPROD(C,Z,M,N,FI,I,FKL,K,L)
C
  DIMENSION C(1),Z(1)
  II=N*(I-1)+I
  KK=N*(K-1)+K
  LL=N*(L-1)+L
  KL=N*(K-1)+L
  ZZ=FKL*(C(KL)-Z(K)*Z(L))
  CC=FKL*FKL*(C(KK)*C(LL)+C(KL)*C(KL))
  G1=FKL*Z(K)
  G2=FKL*Z(L)
  IF (K-I) 11,10,11
10  G=FI+G2
    CALL CADDE(C,Z,M,N,G,I,G1,L)
    GO TO 40
11  IF(L-I) 31,30,31
30  G=FI+G1
    CALL CADDE(C,Z,M,N,G,I,G2,K)
    GO TO 40
31  CALL CADDE (C,Z,M,N,FI,I,G1,L)
    CALL CADDE (C,Z,M,N,1.,I,G2,K)
40  Z(I)=Z(I)+ZZ
    C(II)=C(II)+CC
    RETURN
    END

```

III - SUBROUTINE CRAZE (C,Z,M,N,I,ZM,V)

Arguments d'entrée : C,Z,M,N,I,ZM,V

Arguments de sortie : C,Z

Objet :

Remplace la $I^{i\text{ème}}$ composante de Z par une nouvelle variable, indépendante de toutes les composantes primitives de Z, de valeur moyenne ZM et de variance V.

Listing :

```

SUBROUTINE CRAZE(C,Z,M,N,I,ZM,V)
C
  DIMENSION C(1),Z(1)
  Z(I)=ZM
  IK=I-N
  NI=N*(I-1)
  DO 10 K=1,M
  IK=IK+N
  KI=NI+K
  C(IK)=0.
10  C(KI)=0.
  II=N*(I-1)+I
  C(II)=V
  RETURN
  END

```

IV - SUBROUTINE CESTE (C,Z,M,N,S,AI,I,AJ,J,V,EC,VC)

Arguments d'entrée : C,Z,M,N,S,AI,I,AJ,J,V,

Arguments de sortie : C,Z,EC,VC

Objet :

Remplace l'espérance et la variance de Z par l'espérance et la variance conditionnelle à :

$$S = AI.Z(I) + AJ.Z(J) + E$$

où E est une variable indépendante centrée de variance V.

Le sous-programme calcule également (arguments de sortie EC,VC) l'espérance et la variance conditionnelle à S de E, en application du résultat intermédiaire B (cf. annexe I).

Listing :

```
C      SUPROUTINE CESTE(C,Z,M,N,S,AI,I,AJ,J,V,EC,VC)
      DIMENSION C(1),Z(1),AM(50)
      IK=I-N
      JK=J-N
      DO 10 K=1,M
      IK=IK+N
      JK=JK+N
10    AM(K)=AI*C(IK)+AJ*C(JK)
      AMA=AI*AM(I)+AJ*AM(J)+V
      COR=(S-AI*Z(I)-AJ*Z(J))/AMA
      DO 20 K=1,M
      Z(K)=Z(K)+COR*AM(K)
      AMB=AM(K)/AMA
      DO 20 L=1,M
      KL=N*(L-1)+K
20    C(KL)=C(KL)-AM(L)*AMB
      EC=COR*V
      VC=V-V*V/AMA
      RETURN
      END
```

ORGANIGRAMME D'IDENTIFICATION

D'UN SYSTEME LINEAIRE RECURRENT DU 2^e ORDRE

Nous détaillons ici comment les sous-programmes exposés dans l'annexe II permettent d'établir le programme d'identification pour l'exemple de base du chapitre 3.

Rappelons les équations :

$$y_{t+1} = a_t y_t + b_t y_{t-1} + c_t u_t + \delta_t$$

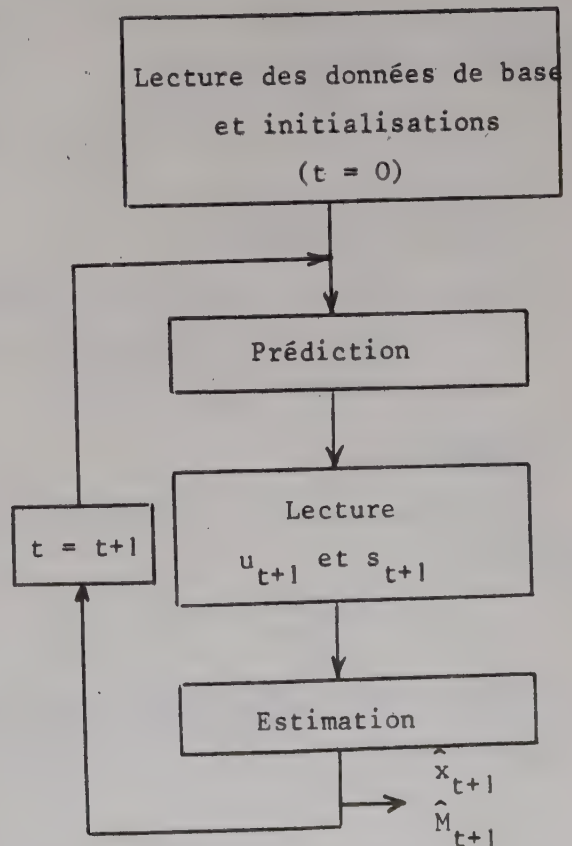
$$s_{t+1} = y_t + \delta'_t$$

Nous avons introduit l'état généralisé :

$$\underline{x}_t = \begin{bmatrix} x_t^1 \\ x_t^2 \\ x_t^3 \\ x_t^4 \\ x_t^5 \\ x_t^6 \end{bmatrix} = \begin{bmatrix} y_t \\ y_{t-1} \\ a_t \\ b_t \\ c_t \\ y_{t-1} \end{bmatrix}$$

(dimension 6)

L'organigramme général sera le suivant :



Reprenons chacune des étapes du programme :

- lecture des données et initialisation

On constitue le vecteur Z des estimations initiales $\hat{x}_0$, accompagné de la matrice des variances covariance $C = \hat{M}_0$.

On introduit la valeur initiale de la commande, soit $U = u_0$.

On définit, d'autre part, les valeurs des variances de δ et δ' , soit, $\text{PHI} = \phi$, $\text{PHIP} = \phi'$.

- étape de prédiction

A ce point du programme, $Z = \hat{x}_t$, $C = \hat{M}_t$. D'autre part, la variable U a la valeur u_t .

Nous effectuons alors la série d'opérations décrite en § 3-1 permettant de passer de $\underline{x}_t$ à $\underline{x}_{t+1}$ en faisant une succession d'appels aux sous-programmes CADDE et CPROD :

↓ composante de Z modifiée par l'instruction

CALL CADDE (C,Z,6,6,0.,6,1.,1)	(remplace y_{t-1} par y_t)
CALL CPROD (C,Z,6,6,0.,1,1.,1,3)	(remplace y_t par $a y_t$)
CALL CPROD (C,Z,6,6,1.,1,1.,2,4)	(ajoute $b y_{t-1}$)
CALL CADDE (C,Z,6,6,1.,1,U,5)	(ajoute $C u_t$)
C(1,1) = C(1,1) + PHI	(ajoute δ')
CALL CADDE (C,Z,6,6,0.,2,1.,6)	(remplace y_{t-1} par y_t)

A la suite de ces instructions, Z et C sont égaux à $\underline{x}_{t+1}^*$ et $\hat{M}_{t+1}^*$.

- lecture des données

On lit : $U = u_{t+1}$ (qui sera utilisé lors du parcours suivant de la boucle)

$S = s_{t+1}$ (qui permet de calculer les estimations $\hat{x}_{t+1}$ et $\hat{M}_{t+1}$)

- estimation

L'information $S = s_{t+1}$ est ici fonction linéaire d'un seul état, mais on peut écrire :

$$s_{t+1} = 1.y_{t+1} + 0.y_t + \delta'_{t+1}$$

d'où l'instruction : CALL CESTE (C,Z,5,6,S,1.,1,0.,2,PHIP,EC,VC)

et alors : $Z = \hat{x}_{t+1}$, $C = \hat{M}_{t+1}$

Remarques :

Il est inutile d'estimer la dernière composante de $\underline{x}$ (égale à y_t) dont nous n'aurons pas besoin par la suite. C'est pourquoi on donne à l'argument M la valeur 5, ce qui permet d'éviter quelques opérations.

En argument de sortie, EC et VC sont les estimations conditionnelles de δ' , grâce auxquelles on peut établir le filtre secondaire d'estimation de la variance de δ' . Par exemple :

$$\text{PHIP} = \text{PHIP} + (\text{GAMMA}/2) * (\text{EC} * \text{EC} + \text{VC} - \text{PHIP})$$

(algorithme simplifié : § 4-3-3)

Conclusion :

Cet exemple montre qu'il n'est pas besoin d'entrer dans le détail des calculs pour établir le filtre. Bien entendu, il serait possible d'élaborer directement le programme d'estimation, pour la classe des systèmes linéaires discrets, par exemple. On ferait alors directement le travail d'analyse correspondant aux équations générales du filtrage non linéaire (§ 2-1), et c'est d'ailleurs ce que nous avons fait primitivement.

Cependant, nos sous-programmes permettent d'adapter rapidement le filtre à beaucoup de situations particulières et tous les sujets que nous avons traité jusqu'ici l'ont été par leur intermédiaire.

B I B L I O G R A P H I E

- AGUILAR-MARTIN J., BANON G.
"Adaptative multivariable algorithm for the stochastic identification of a non linear chemical plant".
IFAC Symp. on Identification, Prague 1970.
- ALBERT E.A., GARDNER A.J.
"Stochastic approximation and non linear regression"
M.I.T. Press, Cambridge, Mass.
- ASTROM K.J., BOHLIN T.
"Numerical identification of linear dynamic systems from normal operating records".
Paper IFAC Symp. Theory of self-adaptative control systems.
Teddington, Engl. 1965.
- ASTROM K.J., EYKHOFF P.
"System Identification"
Survey Paper, IFAC Symp. Prague 1970
Identification and process parameter estimation.
- ATHANS M., WISHNER R.P., BERTOLINI A.
"Suboptimal state estimation for continuous. Time nonlinear systems from discrete noisy measurements".
I.E.E.E. Trans. on Automatic Control. Vol. AC-13, n° 5, oct. 1968
- BARRAUD A., DE LARMINAT Ph.
"Computation of weighting sequences from short duration measurements"
IFAC Symp. Kyoto, 1970.
- BARRAUD A., DE LARMINAT Ph.
"Process identification using its weighting sequence matrix".
IFAC. Symp. Dusseldorf, 1971.
- BRAMMER K.G.
"Imput adaptive Kalman-Bucy Filtering"
I.E.E.E. Trans. on Automatic Control, Fev. 1970, Vol. AC-15, n° 1 pp.157-158.
- BRYSON A.E., FRAZIER M.
"Smoothing for linear and non linear systems"
Proc. Patterson AFB, Ohio TRD-ASD-TRD-63-119, Feb. 1963, pp. 354-364.
- BUCY R.S.
"Non linear filtering theory"
I.E.E.E. Trans. Automatic Control, Vol. AC.10, April 1965.
- BUCY R.S. , JOSEPH P.D.
"Filtering for stochastic processes with application to guidance"
Interscience Publishers, John Wiley and Sons, 1968.

- CHEN R.T.N.
"A recurrence relationship for parameter estimation via method of quasilinearisation and its connection with Kalman filtering"
A.I.A.A. Journal, Vol. 8, n° 9, Sep. 1970.
- CLARKE D.W.
"Generalized-Least squares estimation of the parameters of a dynamic model"
IFAC Symp. Identification in automatic control systems, Prague 1967, paper 3.17
- COLLIN T.
"Emploi des méthodes d'identification dans l'industrie chimique"
D.E.A. d'Automatique, Faculté des Sciences de l'Université de Nantes, Juin 1970.
- COX H.
"on the estimation of state variables and parameters for noisy dynamic systems"
I.E.E.E. Trans. Autom. Control, AC-9, Jan. 1964, pp. 5-12.
- DAWKING J., BRIGGS P.A.N.
"A method of optimal process control using weighing functions as system description"
IFAC. Symp. Teddington, 1965.
- FARISON J.B.
Parameter identification for a class of linear discrete systems"
I.E.E.E. Trans. Autom. Control, AC-12, 1967.
- FAROOK M., MAHALANABIS A.K.
"Application of a statistical approximation method to optimal non linear filtering"
IFAC Symp. on identification and process parameter estimation, Prague 1970.
- HABEGGER L.J., BAILEY R.E.
"Minimum variance estimation of parameters and states in nuclear power systems"
Proc. fourth IFAC Congress, Warsaw, 1969, paper 12.2.
- HOCK F.
"Multivariable and feed forward control by means of self adapting models"
IFAC Symp. on multivariable control systems, Düsseldorf, Oct. 1968.
- JAZWINSKI A.H.
"Filtering for non linear dynamical systems".
I.E.E.E. Trans. on Automatic Control, Oct. 1966.
- KAILATH T.
"An innovation approach to least-squares estimation"
I.E.E.E. Trans. Automatic Control, Vol. AC-13, n° 6, Dec. 1968.
- KALMAN R.E., BUCY R.S.
"New results in linear filtering and prediction theory"
Trans. ASME, J. Basic Eng., serie D, Vol. 83, pp. 95-108, March 1961.
- KOPP R., ORFORD R.J.
"Linear regression to system identification for adaptative systems"
J. of A.I.A.A., Vol. 1, n° 10, pp. 2300-2306, 1962.
- KUSHNER H.J.
"On the differencial equations satisfied by conditional probability density of Markov processes"
J. SIAM on Control, Vol. 2, n° 1, pp. 106-119, 1964.

- KUSHNER H.J.
"Approximation to optimal non linear filters"
I.E.E.E. Trans. Automatic Control, Vol. AC-12, n° 5, Oct. 1967.
- LANDAU I.D.
"A hyperstability criterion for model reference adaptative control systems"
I.E.E.E. Trans. on Automatic Control, AC-14, Oct. 1969.
- LANING H., BATTIN R.
"Random Processes in Automatic Control"
New York, Mc. Graw Hill, 1958.
- DE LARMINAT Ph.
"Commande optimale des systèmes linéaires stochastiques discrets représentés par un modèle "séquence de pondération" - Critère quadratique-Contraintes linéaires".
Annales E.N.S.M., 2ème semestre 1968.
- LEE R.C.K.
"Optimal estimation, identification and control"
Cambridge, Mass. M.I.T. Press, 1964.
- LEVIN M.J.
"Optimum estimation of impulse response in the presence of noise"
IRE Trans. Circuit Theory, CT-7, n° 1, 50-56.
- LOEB J., CAHEN R.
"Identification expérimentale des processus industriels"
Dunod 1967.
- MEHRA R.K.
"On the identification of variances and adaptative Kalman filtering"
I.E.E.E. Trans. on Automatic Control, Vol. AC-15, n° 2, April 1970, pp. 175-184.
- MOWERY W.O.
"Least squares recursive differential correction estimation in non linear problems"
I.E.E.E. Trans. Automatic Control. Vol. AC-10, n° 4, Oct. 1965.
- NEAL S.R.
"Non linear estimation techniques"
I.E.E.E. Trans. on Automatic Control. Dec. 1968.
- PAPOULIS A.
"Probability, Random variables and stochastic processes"
Mac Graw-Hill System Science Series 1965.
- PEARSON J.B.
"A note on non linear filtering"
I.E.E.E. Trans. Automatic Control, Feb. 1968.
- RADIX J.C.
"Introduction au filtrage numérique"
Eyrolles, 1970.
- RICHALET J.
"Methode du modèle"
Rapport CERA, PB.9, Déc. 1966.

- SASTRY V.A., VETTER W.J.
"The application of identification methods to an industrial process"
Preprint JACC, 1969.
- SCHWARTZ L., STEAR E.B.
"A computational comparison of several non linear filters"
I.E.E.E. Trans. Autom. Control AC-13, 83-86.
- SCHWEPPE F.C.
"Evaluation of likelihood functions for Gaussian signals"
I.E.E.E. Trans. Information Theory, Vol. IT 11, pp. 61-70, Jan. 1965.
- STEIGLITZ K., Mc. BRIDE L.E.
"A technique for the identification of linear systems"
I.E.E.E. Trans. on Automatic Control, Oct. 1965.
- STRATONOVITCH R.L.
"Application of the theory of Markov processes to the optimal filtration of signals"
Non linear transformation of stochastic processes.
P.I. Kutnetsov (ed). Pergamon Press, N.Y., 1965.
- TAIT G.R., BELANGER P.R.
"A comparison of some parameter identification schemes using first and second order extended Kalman Bucy filters and sensitivity function"
IFAC Symp. Identification and Process Parameter Estimation, Prague 1970;
- WELLS C.H.
"Application of the extended Kalman estimator to the non linear well stirred reactor problem"
Proc. JACC, 473-482.
- WONHAM W.M.
"Some applications of stochastic differential equations to optimum non linear filtering"
SIAM J. Control, 2, 1965, pp. 347-369.
- ZADEH L.A.
"From circuit theory to system theory"
Proc IRE, 50, 1962, 856-865.

